PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : G01N 33/531, C07K 1/02, 1/10, 3/06, 3/08, 5/02, 7/02 — A1

(11) International Publication Number: WO 95/00846

(43) International Publication Date: 5 January 1995 (05.01.95)

(21) International Application Number: PCT/US94/07222

(22) International Filing Date: 21 June 1994 (21.06.94)

(30) Priority Data:
08/081,412 21 June 1993 (21.06.93) US

(71)(72) Applicant and Inventor: TAM, James, P. [US/US]; 607 S. Wilson Boulevard, Nashville, TN 37215 (US).

(74) Agent: LANQUIST, Edward, D., Jr.; Waddey & Patterson, 27th Floor, L & C Tower, 401 Church, Nashville, TN 37219 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
With amended claims and statement.

BEST AVAILABLE COPY

(54) Title: DOMAIN LIGATION STRATEGY TO ENGINEER PROTEINS WITH UNUSUAL ARCHITECTURES

(57) Abstract

A novel method of chemical ligation of peptide segments requires no protecting groups and no activation of the C-α carboxyl groups. The method consists of four steps: (1) aldehyde initiation in which a masked glycolaldehyde ester is linked to the carboxylic group of an unprotected peptide by enzymatic reaction; (2) ring formation - the regenerated aldehyde reacts with the N-α amine of the second unprotected peptide; (3) rearrangement where O-acyl linkage is transferred to N-acyl linkage to form a peptide bond at higher pH; (4) reconversion to the natural amino acids if necessary. This invention also relates to a method for ligating unprotected peptide, proteins or nonpeptide segments to give therapeutic products and synthetic vaccines with linear, circularized, or branched backbone structures. The method embraces the discovery that when a weakly basic nucleophile and aldehyde are used as a reacting pair for the ligation of two segments, the initial covalent formation of thiazolidine, oxazolidine, hydrozone and oxime bonds is highly specific and may undergo a subsequent intramolecular O to N-acyl rearrangement step which results in the formation of amide bond. Weak bases on a peptide segment are those that contain 1,2- or 1,3-amino thiol or alcohol or those that contain an electronic withdrawing group to the amine such as hydrazide, oxime, phenylhydrazine. Unprotected peptides containing amino terminal residues of cysteine, serine, or threonine can therefore be served as a weak base or as a masked aldehyde which can be generated by oxidation. This invention also relates to a method using the same concept of weak base-aldehyde ligation for site-specific modification of peptides or proteins by lipidation and pegylation. More particularly, the invention relates to the modification of the protein gp120 derived from the human immunodeficiency virus-1 at the amino terminus to contain one or more lipid side chains (lipidation) to increase its efficacy for vaccine and the modificacition of cytokine interleukin-2 by polyethylene glycol (PEG, pegylation) to increase its stability.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

DESCRIPTION

DOMAIN LIGATION STRATEGY TO ENGINEER PROTEINS WITH UNUSUAL ARCHITECTURES

TECHNICAL FIELD

The present invention relates to a method for ligating the carboxylic and the amino groups of one or two peptide segments through an amide bond where the functional groups of the segments are either minimally protected, partially protected, globally protected or not protected at all. More particularly, it relates to a method for ligating one peptide segment to itself or two peptide segments to each other by using a masked aldehyde ester incorporated onto the carboxylic group of a first peptide segment through an enzymatic coupling procedure, which masked aldehyde ester is then released in order that it may react with a ß-functionalized amino group of a second peptide segment to form a ring leading to an O to N-acyl rearrangement step which results in the formation of an amide bond between the peptide segments.

BACKGROUND ART

The synthesis of peptides or proteins has become highly efficient with the advances of the solid-phase peptide synthesis and recombinant DNA technology. Solid-phase peptide synthesis with the aid of automation and other mechanical devices can quickly produce a peptide of greater than 100 amino acids or a library of hundreds of short peptides. The recombinant DNA technology with an optimal expression system can produce proteins accurately and in large quantity. The ideal method of chemical ligation of peptide segments would incorporate both the efficiency of the solid-phase method to generate specific segments and the availability of proteins generated by the recombinant method. The combination of the two types of production of peptide segments would enable engineered proteins to contain unusual structures or nongenetic encoded amino acids by a specific ligation method.

A strong impediment to this approach is a lack of an efficient method for their synthesis. In particular, there is no effective chemical method to selectively couple two unprotected peptide segments to form an amide bond. In general, protecting groups are necessarily attached to nontarget functional groups on the first peptide segment prior

to activation of the C-α of the carboxylic group by a coupling reagent and the consequent peptide bond formation with the N-α of the amino group of the second protected peptide segment. As a result, the developments of the various protecting group schemes have been the key for the conventional approach of ligating peptide segments.

However, the use of protected peptide segments is incompatible with the overall scheme of engineering proteins using proteins produced by means of recombinant DNA technology as a source. It also has limitations of being labor-intensive and unpredictable, partly due to the solubility and coupling difficulties of protected peptide segments. Often, large protected peptide segments are minimally soluble in even the most powerful polar aprotic solvents such as dimethylsulfoxide (DMSO) and dimethylforamide (DMF). The problem of insolubility in protected peptide segments has been addressed with limited success in several ways, including the use of (1) partial protecting group strategy which masks all side chains except those of Ser, Thr, and Tyr, and (2) minimal protecting group strategy which masks only thiol and amino side chains. Protecting groups used in all these approaches alter peptide conformations. This creates a difficult problem in the synthesis of large peptides, since folding and renaturation are required after the completion of the synthesis and removal of protecting groups. These limitations, coupled with the ease of obtaining proteins and protein domains through recombinant DNA technologies, have suggested the need to develop a new strategy for ligating unprotected peptides and proteins in order to engineer new proteins with unusual structures, architectures and functions.

Since protecting groups are the root of the problem, scientists have developed two ligation strategies in the past ten years which use unprotected segments. One of the methods requires the use of enzymes in the reverse proteolysis process in conjunction with a high content of water-miscible solvents. Although enzymatic synthesis has been successful with small peptides, enzymatic synthesis of large peptides has presented difficulties. The stringent criteria demanded by using high molar concentrations of peptide segments accompanied by rapid completion of the reverse proteolytic process without the attendant hydrolysis or transpeptidation have been prohibitive obstacles in the enzymatic synthesis of large peptides. Nevertheless, the use

of enzymes in coupling unprotected peptide segments eliminates the necessity of activating the carboxylic group involved in the coupling reaction of the peptide segments. Furthermore, it also provides the ability to perform the reaction in an aqueous environment.

Another strategy uses a tricyclic aromatic template containing an aryl alcohol and a thiol to form an active ester with the carboxyl segment and a disulfide with the amino segment, respectively, in order to bring two unprotected peptide segments in close proximity with each other. Such positioning of the peptide segments enables them to undergo an O to N-acyl transfer reaction (Fotouhi, N. et al., **1989**; **Kemp, D.S. et al., 1991).**

A problem with the currently accepted methods of protein synthesis which include both conventional liquid state and solid state peptide syntheses is that their application is limited to small straight chain peptide segments, whereas the need exists for such a method of synthesis to be available for long straight chain peptides, branched straight chain peptides and circular peptides.

It is an object of this invention to provide a method of ligation of two peptide segments from the group comprising, but not limited to, long straight chain peptides, branched straight chain peptides and circular peptides, without protecting the various functional groups and without activating the carboxyl group of a first peptide segment which will form a peptide bond with the amino group of a second peptide segment.

In addition, it is an object of this invention to provide a method of ligation of a peptide segment to a compound from the group comprising, but not limited to, DNA by means of incorporating a masked α-aldehyde ester on a carboxylic group and activating that group by releasing the aldehyde thus allowing the carboxylic group to interact with an amino group to form an amide bond. Ligating proteins or peptides to DNA can be useful in biological studies.

It is a further object of this invention that the method developed in this application will make circular proteins readily available for biochemical, biophysical, and therapeutic uses. Another object of this invention is linking multiple copies of unprotected peptides or proteins to a scaffold or template by an amide to produce a branched protein. This application has broad utility. The present method provides a

specific and stable conjugation for peptide/protein antigen to a carrier, drug to a protein, reporter group to an antibody or enzyme, and many others.

Furthermore, it is an object of this invention to provide a high effective molarity for peptide bond formation through the efficient O to N-acyl transfer reaction.

It is an additional object of this invention to provide a versatile means of enzymatic coupling to activate a carboxylic group.

It is also an object of this invention that the reactions required in the method of domain ligation may be run in one vessel in aqueous solution, require only pH changes, no intermediate purification steps and no harsh final deprotection, renaturation or disulfide bond formation.

Finally, it is an object of this invention to provide a means for the site-specific modification of a protein with a peptide, protein, or non-protein molecule such as a lipid or polyethylene glycol.

DISCLOSURE OF THE INVENTION

The present invention relates to a new method to ligate peptide segments chemically without using protecting groups. A major handicap in developing a ligation method using unprotected peptides is that the carboxyl moiety cannot be activated in any form, even as a weak active ester, in the presence of unprotected side chain amines and other nucleophiles. However, if the N-α and C-α peptide segments can be brought close together, their proximity may enable peptide bond formation through an O to N-acyl rearrangement. The domain ligation strategy takes advantage of the O to N-acyl rearrangement reaction. The strategy comprises a series of steps including: (1) aldehyde initiation in which a masked glycolaldehyde ester is linked to the carboxylic group of an unprotected peptide by enzymatic reaction; (2) ring formation in which the regenerated aldehyde reacts with the N-α amine of the second unprotected peptide; (3) rearrangement where the O-acyl linkage is transferred to the N-acyl linkage to form a peptide bond at higher pH; (4) reconversion to the natural amino acids if necessary. Only α-amino acids which have a 1,2-disubstitution pattern and are able to form a five member ring, and 1,3-disubstituted α-amino acids which are able to form a six member ring are able to attain the proper physical spacing which will allow an intramolecular

reaction to occur.

The chemical ligation method known as domain ligation strategy does not require protecting groups or activation of the carboxyl component in the conventional sense.

The domain ligation strategy can be applied to circular proteins as well as branched straight-chain proteins. These proteins are called circular rather than cyclic because they are connected end to end by a peptide bond while cyclization comprises nonspecific circularization. Circular proteins are difficult synthetic targets because several of them contain three disulfide bonds. In addition, they have potential therapeutic values and circularization may improve their half life in vivo and increase their stability against proteolytic degradation, particularly by exopeptidases. Unlike cyclic peptides, synthetic circularized proteins are rare because they are not easily susceptible to synthetic methods presently available. One example of a circular protein is BPTI which uses non-specific carbodiimide for its circularization.

The domain ligation strategy can also be used in the construction of protein dendrimers and in the site-specific modification of proteins because of the specificity of the reaction between the weakly basic nucleophile and the aldehyde in acidic conditions.

Fig. 1 is a chemical formula representation of the general concept of the method of the present invention.

Fig. 2 is a chemical formula representation of an example of the use of the method of the present invention using small compounds.

Fig. 3 is a chemical formula representation of the mechanism of the acyl transfer reaction step of the method of the present invention.

Fig. 4 is a chemical formula representation of an example of the domain ligation method of the present invention, specifically showing the synthesis of a pentadecapeptide--(SEQ ID NO:1)--.

Fig.'s 5A and 5B are HPLC profiles of the rearrangement reaction step of the method of the present invention from compound IId1 to IIe1 as shown on Fig. 2. under the following conditions: A. Reaction at pH 9: a. t = 22 min; b. t = 155 min.; c. t = 284 min.; d. t = 540 min.

B. Reaction at pH 6: a. t = 2.5 h; b. t = 11 h; c. t =26 h; d. t = 61 h;

Peaks 1,2,3 and 4 correspond to compounds IId1, Z-Ala-OH, IIe1, IIc1.

Fig. 6 is an HPLC profile for the segment synthesis of the model pentadecapeptide:

a. before TFA deprotection;

b. after TFA deprotection and upon redissolution in acetate buffer (pH 4);

c. ring formation product after 3 h reaction;

d. purified ring product before rearrangement;

e. 20 h at pH 5; and

f. after 2 days at pH 5.

Fig. 7 is a table showing the rate the acyl transfer reaction which occurs during the method of the present invention.

Fig. 8 is a schematic representation of a peptide dendrimer containing eight peptidyl branches anchored on a scaffolding of oligolysine (indicated by circled K) via a thiazolidine linkage (solid circles) which is obtained by reacting the N-terminal cysteine with a glyoxylyl scaffolding.

Fig. 9 is synthesis of MAPs with formation of 1): oxime-, 2): hydrazone-, and 3): thiazolidine- linkage. Peptide sequence: VA20 = VMEYKARRKRAAIHVMLALA. Reaction product: I): VA20-Oxm-MAP; II): VA20-Hdz-MAP; III): VA20-Thz-MAP.

Fig. 10 is a scheme for preparation of glyoxylyl-MAP core matrix.

Fig. 11 is RP-HPLC analysis of oxime reactions between unprotected peptide and glyoxylyl-MAP [$(CHOCO)_4$-MAP]. Unprotected peptide was marked as 1. MAP cores linked with two, three, or four copies of peptide were labeled as 2, 3, or 4, respectively. HPLC conditions are described in the experimental section.

Fig. 12 is RP-HPLC analysis of hydrazone reactions between unprotected peptide and glyoxylyl-MAP [$(CHOCO)_4$-MAP]. Unprotected peptide derivative was marked as 1. MAP cores liked with two, three, or four copies of peptide were labeled as 2, 3, or 4, respectively. HPLC conditions are described in the experimental section.

Fig. 13 is RP-HPLC analysis of thiazolidine reactions between unprotected peptide and glyoxylyl-MAP [$(CHOCO)_4$-MAP]. Unprotected peptide derivative was

marked as 1. MAP cores linked with three or four copies of peptide were labeled as 3, or 4, respectively. The peak marked with 2' is the dimeric product of peptide derivative formed via disulfide bond. HPLC conditions are described in the experimental section.

Fig. 14 is a set of three tables comparing rates of MAP formation through oxime, hyrazone and thiazolidine. Fig. 15 is MALDI-MS of MAP ligates synthesized through oxime (A), hydrazone (B), and thiazolidine (C) linkage.

Fig. 16 is a general scheme depicting the chemistry of weak base-aldehyde for the ligation of unprotected peptide segments to a MAP core matrix containing an α-oxoacyl group (A) thiazolidine ring formation between a peptide bearing N-terminal cysteine and (B) hydrazone formation between a peptide with N-terminal 4-hydrazino benzoyl group.

Fig. 17 is a synthetic scheme for the preparation of the α-oxoacyl-MAP core matrix.

Fig. 18 is RP-HPLC profile of (A) purified $(CHOCO)_8$-Lys_4-Lys_2-Lys-ßAla-OH using a C18-column; (B) the purified starting material CA16 using a C18-column; (C) the crude conjugation reaction between $(CHOCO)_8$-Lys_4-Lys_2-Lys-ßALA-OH and the peptide Ca-16 using a C8-column; and (D) incubated $(NA\text{-}15)_4$-Thz-MAP at 37 °C,pH 7 after 110 hr using a C8-column to show the stability.

Fig. 19 is RP-HPLC profile of (A) the purified peptide Hob-SR10 using a C18 column; (B) the crude ligation reaction between $(CHOCO)_8$-Lys_4-Lys_2-Lys-ßAla-OH and Hob-SR10 using a C8-column; and (C) incubated $(SR\text{-}10)_4$-Hab-MAP at 37 °C, pH 7.4 after 46.5 hr using a C4-column to show the stability.

Fig. 20 is a table depicting the effects of anti-oxident, organic solvent and temperature on the thiazolidine ring on the ligation of unprotected peptide CA-16 to (CHO)4-MAP.

Fig. 21 is an analysis of peptide fragments and ligation products by matrix-assisted laser desorption mass spectrometry.

Fig. 22 is an ultraviolet spectra of Hob-SR10 (unshaded) and $(SR\text{-}10)_4$-Hab-MAP (shaded).

Fig. 23 is a schematic of the reaction scheme using thiolactones in the site-

specific modification of proteins and a table including examples of substituted thiolactones.

Fig. 24 is a table defining the abbreviated words used in this patent application.

Fig. 25 is a list of Abbreviations.

BEST MODE FOR CARRYING OUT THE INVENTION

The present invention provides a method of chemically ligating a first peptide segment to a second peptide segment. It also provides a method of chemically ligating a peptide to a branched protein or other macromolecule. It further provides a method for the site-specific modification of proteins by chemically ligating such proteins to other peptides, proteins, or non-peptide macromolecules. The method of ligation involves reacting a weak base with an aldehyde to form a bonding structure that can undergo a subsequent intramolecular O to N-acyl rearrangement that results in the formation of an amide bond. Weak bases are any nucleophiles that are not protonated at aqueous acid pH and can react with an aldehyde at such pH. These may include 1,2-substituted aminothiol such as cysteine, 1,2-substituted aminoethanol such as serine and threonine as well as those amines which are substituted at the α position with electronic withdrawing groups e.g. heteroatoms such as oxime, substituted hydrazines.

The method of ligation two peptide segments has four basic steps: (1) aldehyde initiation in which a masked glycolaldehyde ester is linked to a carboxyl group of a peptide segment by enzymatic reaction; (2) ring formation in which the regenerated aldehyde reacts with the N-α amine of the second unprotected peptide containing either Cys or Thr to form a thiazolidine or oxazolidine ring, respectively, at relatively acidic pH; (3) rearrangement through an O to N-acyl transfer reaction; (4) reconversion to natural amino acid residues if necessary.

The method of domain ligation brings two unprotected peptides together with unusual regiospecificity through their respective carboxyl and amino functional groups. To achieve this, simple alkyl aldehydes are introduced at the carboxyl terminus of one peptide segment that will then react selectively with the N-α amino group of the second peptide component. The two peptide segments which are to be ligated could comprise opposite ends of the same polypeptide segment, if circularization of the peptide is desired. Ligation of two separate peptide segments will result in a straight chain peptide segment.

Aldehydes condense with amines to form imines which are unstable and reversible in aqueous solution unless a ring or conjugated system is formed. The reversibility of the reaction of aldehydes with amines and the stability of ring formation

with ß-functionalized amines is crucial in the design of this chemical ligation strategy. Side chain amines which do not contain a ß-functionalized thiol or hydroxyl group are incapable of ring formation. Specifically, an alkyl aldehyde is introduced via an ester linkage to a carboxyl group of one peptide segment to be captured by the second peptide segment bearing a ß-functionalized amino-terminal amino group (such as those occurring in Cys, Thr and Ser residues) to form a relatively stable five or six-member ring (Figure 1). Only α-amino acids which have the 1,2-disubsitution pattern or the 1,3-disubstitution pattern will allow ring formation. Side chain functional groups are not capable of forming the ring. Furthermore, this reaction is usually performed at pH 5 or 6 to further avoid Schiff base formation with side chains of Lys or Arg. The net result is that the carboxyl and amino components are brought together by a ring formation leading to a well positioned and facile intramolecular O to N-acyl rearrangement to form the desired amide bond. Thus, peptide bond formation occurs without activation by a coupling reagent which is an invariable feature of the conventional approach.

The two reacting termini could be from the same peptide if the intention is for circularization. The ester bond which links the peptide segment and the aldehyde is positioned in such a way that an amide bond can then be formed through an intramolecular O to N-acyl transfer reaction (Figure 1). The domain ligation strategy employs the α-acyloxyacetaldehyde system (a glycolaldehyde ester, or an α-formylmethyl ester of the carboxylic component). In this system the acyl carbonyl and the aldehyde carbonyl are separated by two atoms to facilitate a 5-member ring or by three atoms to facilitate a 6-member ring transition state.

EXAMPLE 1

The initial step of this example involves the reaction of Z-Ala with a masked acetal containing the aldehyde function (formylmethyl ester, FM) as the carboxyl component which was obtained by reacting the Cs salt of Z-Ala with bromoacetaldehyde dimethyl acetal in DMF. Acetal is particularly suitable because of its ease of removal and it gives an unreactive side product (MeOH) which eliminates the need for a purification process. Treatment with 30% TFA in acetonitrile or the most TFA-compatible organic solvent used in peptide synthesis for acid deprotection,

such as $CHCl_3$, CH_2Cl_2, CCl_4, toluene, etc., in a trace amount of H_2O at 0° C for 15 to 20 min smoothly converted the acetal to its aldehyde Z-Ala-OFM. After TFA and the solvents were removed, the aldehyde was allowed to react without purification with ß-mercaptoethylamine, cysteine, cysteine methyl ester, or threonine methyl ester to form thiazolidine or oxazolidine, which rearranged to the amide form at higher pH.

The reaction of carbonyl compounds with cysteine involves an initial reaction of sulfhydryl with the carbonyl to give an addition product which condenses with the N-α of the amino group of the cysteine to form a cyclic thiazolidine derivative over a wide range of pH. To avoid the hydrolysis of the ester and the unwanted reaction of aldehyde with the side chain amino groups, the reaction of cysteine or other ß-mercaptoamines was conducted with aldehydes at pH 4 to 5. Under these conditions the thiazolidine product was formed almost immediately. At lower pH, this reaction was slower. At pH 2 the reaction required 1-2 hr for completion. The thiazolidines were stable and were easily purified by normal or reversed phase HPLC under usual conditions.

EXAMPLE 2

The basic character of the secondary amine in the thiazolidine ring makes it possible for the acyl group to migrate from the ester oxygen to the nitrogen. The O to N-acyl transfer reaction is a dominating side reaction in the acidic deprotection step of peptide synthesis during which the acyl moiety of the peptide migrates from the α-amine to the free hydroxy group on the side chain of a serine or threonine residue. The transfer reaction is reversible upon base treatment involving a 5 member ring oxazolidine-like transition state. Figure 7 shows that O to N-acyl rearrangement was effective even at the acidic pH range. The weak basicity of the thiazolidine amine (pKa 6.2) may have contributed to this. Rearrangement occurred in all the pH conditions from pH 6 to pH 9. The rate of the reaction was largely dependent on the pH value. As shown in Figure 7, for IId1 to IIe1, at pH 9, the $t^{1/2}$ for the rearrangement is about 3.4 hr and the reaction proceeds cleanly without any detectable side products, while at pH 6 the $t^{1/2}$ is about one day with some side products probably due to the regenerated aldehyde by slow reversible hydrolysis.

EXAMPLE 3

Both the condensation products and the rearranged products give a mixture of two diastereoisomers due to the creation of a new asymmetric carbon at position 2 of the thiazolidine ring. These diastereoisomers are HPLC separable. The difference between the ester (before rearrangement) and amide (after rearrangement) forms is distinguishable in the following ways: analytically (HPLC), spectrometrically (NMR) and chemically. In the NMR studies, the prominent changes are the disappearance of the proton signals for the secondary ammonium protons in the thiazolidine ring and an up-field shift of two protons on the methylene carbon linked to the oxycarbonyl which, after rearrangement, became an hydroxyl group. The ester form is susceptible to alkali hydrolysis under saponification conditions, whereas the amide form is stable. Treatment of the ester product IId1 with 0.1 to 1 N NaOH gave Z-Ala-OH as the hydrolyzed product along with the rearrangement product in 10 min, while the amide form IIe1 was stable under the same conditions.

The rearrangement product now resembles a Z-Ala-Pro structure with a thiol ether linkage as the isoelectronic replacement of the methylene carbon at position 4 and an hydroxymethyl substitution on position 5 of the proline ring. Such modifications will unlikely change the backbone conformation of a proline-containing peptide chain. Thus, this pseudo X-Pro bond can generally be viewed as a substituent for any of the X-Pro bonds present in protein sequences and further enlarges the scope of application of the domain ligation strategy. The N-acyl thiazolidine structure is similar in structure to an acetamidomethyl protected form (Acm-like) of the cysteine residue, and can be reverted to the thiol to form a disulfide by the usual oxidative cleavage method.

EXAMPLE 4

The activation step whereby the formylmethyl ester is introduced to the unprotected peptide segment is an essential element of this method of peptide ligation. With synthetic peptides, this should not pose a problem, since new resins have been developed to give such a linker functional group containing a formylmethyl ester at the carboxyl moiety. For proteins derived from recombinant DNA or natural sources, an activation step is needed. The key is the introduction of the masked aldehyde function onto the carboxylic group of the first component by using the specificity of an enzyme.

Kinetically controlled aminolysis by enzymes of a peptide ester in the presence of a water miscible organic solvent could be smoothly and efficiently accomplished if a high concentration of the amino component is used. In this case, a large excess of dimethoxyethyl ester of alanine was used. The carboxylic component peptide ester was synthesized by solid phase peptide synthesis method using a newly developed resin (see Fig. 4: IVa). After the cleavage and removal of all protecting groups, this resin provided a peptide containing a carboxyl ester (Fig. 4:IVb)--(SEQ ID NO:3)-- which is a good substrate for enzyme catalyzed coupling. The enzymatic coupling between the peptide and the small substrate, the 2,2-dimethoxyethyl ester of alanine, abbreviated AlaODMoE (IVc) was catalyzed by trypsin in 60% or more DMF and completed within a short time (< 1/2 hr) under the condition of high molar concentrations and large excess of the amino component (Fig. 4:IVc). After the enzymatic coupling, the capture and all subsequent steps were carried out in a similar way as described in the model study. The hexapeptide (Fig. 4:IVe) (SEQ ID NO:5) was premixed with the first peptide acetal (Fig. 4:IVd) (SEQ ID NO:4). The deprotection was achieved using 95% TFA containing 5% H_2O for 3-5 min at 0°C. The regenerated aldehyde showed a shorter retention time in reverse phase HPLC. The HPLC did not reveal any detectable hydrolysis of the ester bond. After removal of TFA, the reactants were redissolved in an acetate buffer (pH 4) to allow the capture reaction, which was also monitored by HPLC. The strong affinity between the aldehyde and the ß-thiol amino compound makes it possible that the components can be used in a very dilute molar concentration. This is highly desirable since the reaction between macromolecules has to be carried out in a dilute solution due to the large molecular masses of the reactants. Under the present experimental conditions, the concentration of both components is about 5 mM. The expected thiazolidine product formed cleanly without any detectable side reactions. The rearrangement reaction occurred after the solution was adjusted to pH 5 with aqueous acetate buffer.

All the side chain functional groups in both peptides were unprotected, including the ϵ-amine of lysine and the ß-carboxylic acid of aspartic acid residues, which inevitably have to be protected in the conventional segment coupling approach. It is also important to note that the three steps of this method (acetal deprotection, aldehyde

capture and acyl transfer) could be conducted in the same reaction vessel, required only pH changes in aqueous solution and no intermediate purification steps were needed. The product formed (Fig. 4:IVg) (SEQ ID NO:1) also did not require renaturation and oxidation to form the disulfide bond. This has simplified the experimental procedure significantly.

The domain ligation strategy for forming a peptide bond between the opposite ends of one unprotected peptide segment or two separate unprotected peptide segments has high selectivity and efficiency and reaction rate, as well as the ability to be conducted under aqueous conditions, all of which are essential elements of a useful chemical ligation method. The domain ligation strategy meets all of these requirements. The unique reactivity of an aldehyde with a ß-mercaptoamino compound under acidic conditions makes it unnecessary to protect other functional groups. This reactivity also accomplishes the ligation of two components with a high effective molarity through the efficient O to N-acyl transfer reaction. Furthermore, this rearrangement permits the formation of an amide bond without the participation of an activated carboxylic group.

The N to O-acyl transfer reaction is a common side reaction in peptide synthesis and usually occurs in anhydrous acid treatments to give an ester linkage similar to that formed in domain ligation strategy. This side reaction is reverted back, i.e. the reverse O to N-acyl transfer reaction to give the normal peptide bond, with great efficiency using a base treatment at pH 8-9. The modified Thr (oxazolidine) and Cys (thiazolidine) can be converted to their respective amino acid by treatment with aqueous base.

Some therapeutic applications of this method include production of proteins having unusual architectures. Selected proteins for design are circularized and branched proteins. Target circularized proteins include interleukin-1 receptor antagonist which is currently in clinical trials as a drug to reduce severity of sepsis and arthritis; monitor peptide which is a cholecystokinin-releasing factor and may be useful for treatment of digestive disorders; and defensin which is a broad-spectrum antibiotic with promising activity against AIDS-related pathogens. Target branch proteins will include a malaria vaccine containing the protective antigen derived from merozoite surface protein (MSP-1). This antigen is the most promising vaccine candidate to date.

Domain ligation strategy is a method to link or circularize totally unprotected peptide and protein segments via a peptide bond without activation. This method is well suited for the synthesis of circularized and branched proteins which are inaccessible directly by recombinant DNA methods and are difficult to obtain by the conventional methods of peptide synthesis. The domain ligation strategy employs a combined approach of organic and peptide chemistry in engineering proteins for therapeutic applications.

Circular proteins differ from cyclic proteins because they are connected end-to-end by a peptide bond, while cyclization is a non-specific circularization. These cyclized proteins represent difficult synthetic targets because several of them contain three disulfide bonds. However, methods developed for their synthesis would be applicable to most other peptides and proteins. In addition, they have potential therapeutic values and circularization may improve their half-life in vivo and increase stability against proteolytic degradation, particularly the exopeptidases. Unlike cyclic peptides, synthetic circularized proteins are rare because they are not easily accessible to methods presently available. One example, BPTI, uses non-specific carbodiimide for its circularization. Thus, methods developed in this application will make these circular proteins readily available for biochemical, biophysical, and therapeutic evaluations. Furthermore, the domain ligation strategy can be extended and amplified to other applications, such as ligating proteins or peptides with DNA for biological studies.

Another application of the domain ligation strategy is linking multiple copies of unprotected peptides or proteins to a scaffold or template by an amide to produce a branched protein. This application has broad utility and this method would provide a specific and stable conjugation for peptide/protein antigen to a carrier, drug to a protein, reporter group to an antibody or enzyme, and many others.

EXAMPLE 5

A method to study this reaction with all possible combinations of N-terminal amino acids and their side chain functionalities using a library of 400 dipeptides consisting of 20 genetically coded amino acids was employed. To this end, identical

copies of a library of 400 dipeptides anchored on cellulose paper were synthesized and an alanyl ester aldehyde was allowed to react with each library under various conditions. The paper support used in these experiments served both as the solid support on which the 400 peptides were synthesized and as a monitoring device on which the ring formation could be observed.

The library was synthesized on Whatman paper using the Fmoc chemistry. The arrangement of the library was in a matrix system containing 400 spots, each representing a dipeptide. The reactivity of each dipeptide towards the aldehyde could be visualized as either a horizontal row which showed the N-terminal amino acid was reactive (e.g. Cys-X, where X represents 20 amino acids) or a vertical column which indicated that the side chains of the carboxyl terminal amino acids (e.g. X-Cys) were reactive. Furthermore, the subsequent O to N-acyl rearrangement could also be observed efficiently by changing to basic buffers or solvents under which the ester bond will be hydrolyzed and the O to N-acyl transferred product would be stable.

For the library to work efficiently, a reporter molecule on the amino acid alkyl ester aldedyde would be required to provide detection of the bimolecular reaction between the aldehyde and the dipeptides. We selected dye-labelling as a reporter group because of its high sensitivity for visible detection, stability under normal conditions, and ease of attachment to the amino acids or peptides. The color functional molecule, 2,4-(dimethylamino)phenylazobenzoic acid (Methy Red, Dpab for abbreviation) was introduced to an amino acid derivative through an amide bond which is stable toward acid and base treatments. Depending on the pH of the aqueous buffer, Dpab-amino acids possess intensive orange to red color and are visible on the paper at low concentration.

Three different aldehydes were used and were esterified to Z-Ala with the following alcohols (1) α-formylmethyl (FM) alcohol, (2) ß-formylethyl (FE) alcohol, and (3) ß,ß,ß-dimethylformylethyl (DFE) alcohol. The FM ester aldehyde will give a five-member transition state and should be 30 to 100 fold faster than either FE or DFE esters which requires a 6-member transition state in the O to N-acyl transfer reaction. ME ester rearranged approximately 100 fold faster than the hindered DFE ester and 25 fold faster than FE ester. ME ester is the ester of choice.

Six different amino acids (Cys, Thr, Ser, Trp, His and Asn) are known to form ring products with simple alkyl aldehydes, in particular with formaldehyde which has been used industrially for tanning and medically for inactivation of toxins and biological agents. Because the ME ester aldehydes could be viewed as simple alkyl aldehydes, ring formation is expected to a certain extent with these six amino acids.

Dipeptides with N-terminal Cys, Thr, and Ser (Cys-X, Thr-X, and Ser-X; where X is any amino acid) are of major interest because their ring products can be reverted to Cys, Thr, or Ser. Thiazolidine and oxazolidine can be viewed as temporary protecting groups for these amino acids. Further, the relatively common occurrence of these amino acids in proteins makes them convenient points for ligation in our strategy. The two heterocyclic amino acids Trp and His bearing weakly acidic amines are known to react with alkyl aldehydes to form bicyclic compounds. The initial kinetic ring product with the heterocyclic amine would further rearrange to the stable product involving the C-2 carbon and N-α of the amino group. These bicyclic ring products are not reversible to their natural amino acids. Similarly, the side chain amide of Asn is also known to participate in ring formation, usually under forcing conditions.

EXAMPLE 6

Ring formation in both aqueous buffered solutions at pH 5 to 8 and 90% water-miscible organic solvents at pH 7 has been studied. However, the order of reactivity was significantly different in both systems. N-terminal Cys reacted rapidly and completely with all three aldehydes at all pH ranges tested. With the unhindered FM and FE esters, the reaction was completed within 0.5 hr but required 2 hr for completion with the hindered DFE ester. In contrast, the reactivity of Thr was 500 fold slower. Ser was basically non-reactive. In general, the reactivity of the N-terminal amino acids with Dpab-Ala-O-FM could be divided into three groups. First, Cys-X reacted exceptionally fast with Dpab-Ala-OFM in either aqueous or water-miscible organic and water mixtures. The reactions were completed in 0.5 h at pH 5-8 even at a very dilute concentration of 1×10^{-6} M. Second, Thr-X, Trp-X and His-X represented a category that reacted 500 to 100,000 fold slower than Cys-X. Their reactivity was highly dependent on concentrations of Dpab-Ala-O-FM, pH, and the neighboring amino acid. Trp-X formed a heterocylic compound in 30-50% in 160

hr at the acidic pH. Furthermore, Trp-X reacted faster than Thr-X when the concentration of Dpab-Ala-O-FM was lower than 5 x 10^{-5} M probably due to the irreversibility of the Trp-X product.

In contrast, the oxazolidine ring and the Schiff base of Thr-X were not stable at the acidic range and only 5-30% of oxazolidines could be observed in 160 hr. At neutral and basic pH, Thr-X reacted faster than Trp-X and 20-60% of oxazolidines could be observed. N-terminal His formed a heterocyclic compound but its formation was slow at pH 5-8 and less than 10% of product was observable in 160 hr. Third, the ring formation with Ser-X and Asn-X was essentially insignificant in aqueous solution. Asn-X reacted very slowly to form the heterocyclic compound. However, in aqueous solution less than 5% of the reaction product could be observed in 160 hr.

ß-carboxamide participates in ring formation with alkyl aldehyde. Primary amides such as Asn, Leu-NH_2, and AlaNH_2 would react with Z-Ala-O-FM to form a heterocyclic compound. Interestingly, Gln-X which would have formed a six-member ring did not react with Dpab-Ala-O-FM probably due to the slow formation of the six-member ring. Similarly, the reaction with Ser was much slower than Thr and never went to significant completion in the aqueous condition because the opened form of Schiff base and hydrolysis were favored.

The neighboring amino acids exert either rate enhancement or retardation in ring formation. When the neighboring group amino acid is hydrophobic such as X-Ile, X-Phe, X-Trp, X-Leu, X-Val, and X-Tyr, ring formation was accelerated when compared with X-Ala and X-Gly. This was particularly evident with those dipeptides containing N-terminal amino acids such as His and Ser which exhibited slow ring formation. In contrast, when the neighboring amino acids are hydrophilic and particularly acidic, such as Asp, Glu and Asn, ring formation is retarded. A possible explanation for the observed result might be the participation of the side chains in assisting the hydrolysis of the ring form to the open form or the Schiff base to the starting material. The rate enhancement of the neighboring amino acids might be due to the hydrophobic interaction of the Dpab which contains two phenyl rings with the hydrophobic sequences.

Except for Cys-X, ring formation in 100% aqueous solutions and in a very

dilute concentration of Dpab-Ala-O-FM was slow for Thr-X and Trp-X, and insignificant for His-X, Ser-X and Asn-X. The equilibrium favors the open forms of either the hydroxymethyl derivatives or the Schiff base which forms are hydrolyzed by water to the starting materials. However, the equilibrium would be predicted to favor the closed forms in the absence of water. To accelerate ring formation, we experimented with the use of 90% water-miscible organic solvents such as hindered alcohol (isopropanol) and aprotic polar solvents (DMF and DMSO). The use of these water-miscible organic solvents in high concentrations are necessary for reverse proteolysis and compatible with the scheme of using unprotected peptide segments. Furthermore, water-miscible organic solvent and water have been applied to effect the incorporation of the amino ester aldehyde in the first step of the domain ligation strategy.

When the reaction was performed in 90% water-miscible organic solvent buffered to pH 7 in 10% H_2O, alcoholic solvents such as isopropanol did not improve the reaction rates. In contrast, the polar aprotic solvents DMF and DMSO greatly accelerated the reaction rates, particularly for Thr-X (except when X is Asp, Glu, or Asn) to give the oxazolidine ring formation in 20 hr (Fig.1). The rate acceleration for Ser-X was difficult to quantify since only about 25% of oxazolidine was observable in 30 hr. Nevertheless, it represents an increase of about 10 fold when compared to the 100% aqueous solution. The rate enhancement was also found in the ring formation of Trp-X, His-X and Asn-X, but apparently the formation of thiazolidine and oxazolidine was more favored.

After comparing the steric effect of the carboxyl terminus bearing the OME ester aldehyde it was found that the rates of O to N-acyl transfer reaction greater favors small amino acids with Gly > > Ala > Val. Gly is the preferred C-terminus residue.

When preparing proteins containing multiple disulfide bonds, the stability of the ester during the folding, renaturation and disulfide formation, which are usually performed under basic pH and requires long duration of 2 to 3 days, is a major concern. Under such conditions, the hydrolysis of the ester is likely to occur and would lower the yield of the subsequent reaction. For this reason, a new method of peptide ligation has been developed that allows the renaturation, and selective disulfide

bond formation for peptides and proteins. The key reagent in this new method was dimethylsulfoxide (DMSO). Facile disulfide bond formation by DMSO in aqueous buffered solutions proceeded across a wide range of pH, from acidic to basic. This result will allow the folding and disulfide oxidation to be performed over an acidic range of 5-6 or near neutral pH of 6-7 where hydrolysis will not be significant. The DMSO oxidation overcame the limitation of the conventional oxidation method using air or mixed disulfide that was applicable only over a narrow basic pH range. The sulfur-sulfur bond reaction by DMSO was selective and no side reactions were observed with nucleophilic amino acids such as Met, Trp, or Tyr. Detailed kinetic studies on a series of monocyclic agonist peptides of ßFGF showed that disulfide formation by 20% DMSO was completed in 0.5 to 4 hr, while similar experiments by air oxidation at basic pH required longer duration and produced incomplete reactions. Facile oxidations by DMSO were observed with the basic and hydrophobic, tricyclic 29-residue human defensin and 78-residue heparin-binding EGF. In contrast, air oxidation at basic pH of these molecules led to extensive precipitation and low yields. DMSO is a versatile and useful oxidizing agent for peptides over a wide range of pH and may be particularly suitable for renaturation and oxidation of proteins at acidic pH of the domain ligation strategy.

EXAMPLE 7

TGFα (transforming growth factor-α) is a 50-residue, three-disulfide protein. It contains two domains which are stable structures and we have shown that they can be folded to give the correct disulfide structure. We synthesized two subdomains TGFα1-32 (SEQ ID NO:6) and TGFα33-50 (SEQ ID NO:7) on a new resin (e.g. hydroxyethyloxymethyl resin) by the solid-phase method, refolded, and purified to give an carboxyl-ester and were ligated chemically as shown in Fig.3. The resulting Thr-33TGFα (SEQ ID NO:2) has the correct molecular weight and the biological activity comparable to TGFα. This synthesis validates the concept of domain ligation strategy and points to its potential of preparing proteins with unusual structures that may not be accessible from recombinant products.

Examples 8 to 10 illustrate the facile and specific nature of the domain ligation to form peptide dendrimers. Peptide dendrimers with their characteristic branched

structures represent a class of artificial proteins assembled on a scaffolding or template and which would attain the macromolecular bulk as proteins, but have the advantages that they self-assemble and obviate the need of extensive folding required for biochemical activity. The flexibility of designing scaffoldings and the attendant dendritic peptides has led to successful engineering of artificial proteins which function as enzymes, ion channels, antibiotics, diagnostic reagents, and vaccines.

Current methods of stepwise solid-phase synthesis of peptide dendrimers are inadequate to yield such macromolecular products with high purity. Although the use of protected peptide monomers offers improvements, it suffers the limitations of poor solubility and slow coupling reactions. A more direct and efficient approach is the use of nonpeptidyl linkages for the ligation reaction between the unprotected peptide segments and scaffoldings. Examples of this approach include conjugation through thioalkyation, thioester, and oxime. Other applicable but yet untried methods include hydrazone, reverse proteolysis, and domain ligation.

Domain ligation is particularly appealing because it utilizes the facile and chemoselective reaction between a weak base such as 1,2-amino thiol and an alkyl aldehyde to give thiazolidine ring under acidic condition. With small peptides, this reaction is usually completed within 10 min and highly specific for the N-terminal of cysteine. Unprotected side chains of lysine, arginine and other amino acids are excluded from this reaction to allow totally unprotected peptides to be ligated to an aldehyde-containing scaffolding.

Example 8 illustrates the utility of the domain ligation in the synthesis of peptide dendrimers for synthetic vaccine purposes, using an octavalent lysinyl scaffolding, popularly known as MAP (multiple antigen peptide) which consists of several levels of sequentially branched lysine.

Example 9 compares three different methods of performing the ligation using thiazolidine, hydrazone, and oxime, in the preparation of another synthetic vaccine for feline leukemia virus using the MAP approach. In each case a homogeneous compound was obtained.

Example 10 illustrates two different methods for the preparation of synthetic vaccine derived from HIV-1.

Example 8

This example addresses the synthesis of a synthetic vaccine comprising an octameric peptide dendrimer consisting of the V3-loop of gp120 of HIV-1, MN strain.

The core matrix, consisting of a masked aldehyde, (Ser_8Lys_4-Lys_2-Lys-β-Ala), was prepared by the solid phase peptide synthesis using the Boc-chemistry, cleaved from the resin support, and purified by reverse-phase HPLC to homogeneity. The alkyl aldehyde in the scaffolding was generated by oxidizing the 1,2-amino ethanol moiety of the N-terminal Ser on the scaffolding, [$Ser_8Lys_4Lys_2$-Lys-β-Ala (Ser_8-MAP)], with sodium periodate at pH 7 to yield a glyoxylyl derivative of $(HCO)_8$-MAP in nearly quantitative yield.

The monomer containing the weak base to be reacted with the aldehyde contained the 1,2-amino thiol group of cysteine at the N-terminal of a purified and unprotected 24-amino acid residue peptide, CI-24 (CNYNKRKRIHIGPGRAFYTTKNII), obtained by the solid phase method. Rusche, J.R., Javaherian, K., McDanal, Pedro, J., Lynn, D.L., Grimaila, R., Langlois, A., Gallo, R.C., Arthur, L.O., Fischinger, P.J., Bolognesi, D.P., Putney, S.D. and Matthews, T.J. (1988) *Proc. Natl. Acad. Aci. U.S.A.* **85**, 3198-3202; Goudsmit, J., Debouck, C., Meloen, R.H., Smith, L., Bakker, M., Asher, D.M., Wolff, A.F., Gibbs Jr, C.J. and Gajdusek, D.C. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4478-4482; Javaherian, K., Langlois, A.J., McDanal, C., Ross, K.L., Eckler, L.I., Jellis, C.L., Profy, A.T., Rusche, J.R., Bolognesi, D.P., Putney, S.D. and Matthews, T.J. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6768-72; Devash, Y., Calvelli, T.A., Wood, D.G., Reagan, K.J. and Rubinstein, A. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 3445-49; Takahashi, H., Cohen, J., Hosmalin, A., Cease, K.B., Houghten, R., Cornette, J.L., DeLisi, C., Moss, B., Germain, R.N. and Berzofsky, J.A. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 3105-3109. Laser-desorption-mass spectrometric analysis of CI-24 (calcd) 2864.4, (found) 2868 and amino acid analyses of these compounds agreed with the calculated values.

The thiazolidine ligation reaction was adequately performed at pH 4.2 in H_2O. However, the use of an organic cosolvent and elevated temperature (37°C) provided consistently better results than in H_2O alone because they enhanced the rate of

formation and prevented various intermediate dendrimers aggregating or precipitating during the course of the reaction. The best combination was found to be N-methylpyrrolidinone (NMP):H_2O, (1:1, v/v). Other organic cosolvents such as DMSO or DMF were not suitable. DMSO was shown to be a mild oxidant that led to disulfide formation of cysteinyl containing peptides and DMF led to formylation (M+28) of the unprotected peptide as shown by MS analysis of products containing M+28 peaks. Using the optimized NMP:H_2O mixture, the less-hindered tetra- and pentameric MAPs were completed in < 2 hr, while the hexa- and heptameric MAP required 7 and 30 hr respectively. The more hindered, fully substituted octameric MAP was found to give 82% in 67 hr. (For a schematic of this reaction see figure 8).

The procedure used for thiazolidine ring formation was as follows. All solutions used for this reaction were deaerated and purged with argon, and EDTA (O.02M) was added to prevent disulfide oxidation. Peptide CI-24 (7.18 mg) dissolved in 0.02M NaOAc (0.13 mL) containing 0.02M EDTA to give a 20 mM solution at pH 4 were added to a 11.6mM solution of MAP core (0.166 mg) in H_2O (0.01 mL) and diluted with an equal volume of NMP:HOAc (9:1, v/v) to give a final volume of NMP:H_2O at 1:1. The reaction was performed at 37°C for 80 hr.

Example 9

A comparison of different ligation methods using thiazolidine, hydrazone, and oxime and a 20-residue antigenic peptide was performed (See Figure 9 for a schematic of reactions).

A 20-amino acid-residue antigenic peptide derived from the surface protein of feline leukemia virus, VA-20 (VMEYKARRKRAAIHVMLALA), containing a weak base such as aminooxy acetyl, mono-hydrazide succinyl, or cysteine attached at the amino terminus was synthesized by stepwise solid phase peptide synthesis employing Fmoc-chemistry on p-benzyloxybenzyl alcohol resin (Wang-resin). After completing the peptide sequence assembly, the functional residues for ligation, aminooxyacetyl, mono-hydrazide succinyl, and cysteinyl were introduced respectively onto the peptide N- termini. Cleavage of the peptides was carried out with TFA11 and crude peptide products were purified by RP-HPLC. The purified peptides were characterized by amino acid analysis and Matrix Assisted Laser Desorption Ionization-MS

(MALDI-MS).

Ligation of peptides with MAP core matrix (See Figure 10 for a schematic of synthesis of matrix.): Typically, the ligation reaction was performed in aqueous condition and this was used as the starting point for later optimization. This "standard condition" used 2.5 molar equivalents of unprotected peptide with concentrations of 1.25 mM and 0.125 mM of MAP core in the aqueous medium buffered by 0.1 M Na/HOAc. The pH ranges of the reaction solution for each type of ligation were selected according to prior knowledge of the reaction mechanism. The pH 4.7, 5.2, and 4.5, therefore, were used for oxime, hydrazone, and thiazolidine ligation, respectively. The reaction process was followed by HPLC analysis (Fig. 11-13). In figures 11-13, all the peptide monomers were marked as peak 1. In oxime and hydrazone ligation reactions, ligation intermediates with one copy of peptide attached on the MAP core were not observed and only MAP core matrix with two (peak 2), three (peak 3), and four (peak 4) copies of peptide were observed during the course of the reaction. In thiazolidine ligation (Fig. 13) only products containing three (peak 3) and four copies of peptide (peak 4) were observed after mixing the two components. The minor peak marked with 2' is a dimer of the peptide linked with a disulfide bond. In the thiazolidine formation, 0.01 M of EDTA was added to the reaction mixture with argon purging through the reaction solution to suppress disulfide formation. Under such a reaction condition, less than 2% of total starting cysteinyl- peptide was oxidized to disulfide dimer (Figure 13, peak 2'). This result shows a significant advantage of ligation under such an acidic condition over other conjugations employing thiol functionality under neutral or basic conditions, which have disulfide formation as one of the major side reactions.

Under an acidic condition, all three reactions produced clean ligation products in good yield, which fulfilled the objective of overcoming side reactions that occur in basic medium. The final ligation products as well as the ligation intermediates were characterized by MALDI-MS. (Fig. 15) Amino acid analysis results of the final ligation products also gave the expected composition. The results of thiazolidine (pH 4.5), oxime (pH 4.7), and hydrazone (pH 5.2) formation in aqueous as well as water-miscible organic solvent-water mixture solution are summarized in Tables 1-3 in

Figure 14.

In general, the ligation reaction rates were increased by manipulating pH, temperature, and organic cosolvents. The optimal reaction pH range for oxime and hydrazone ligations is around 5.0 and 4.5 for thiazolidine ligation. At 37°C the reaction led a 2-3 fold of rate increase. Addition of organic cosolvents further accelerates the reaction rates. DMSO is the most useful cosolvent for oxime and hydrazone formation, while DMF is best for thiazolidine formation. When reactions were performed at 37°C in media containing the appropriate organic solvents, a greater than 10 fold rate increases were observed for all three reactions. Ligation reactions were completed within 6 h, compared to 1-3 days under unoptimized conditions. The thiazolidine ligation showed the fastest reaction rate among the three types of reactions tested and its ligation product was stable in a wide pH range of 3-9. The side reaction of thiol group oxidation can be suppressed by carrying out ligation at a pH as low as 4.5 and by adding EDTA to the reaction medium. The facile reaction rate and stable product make thiazolidine ligation the most attractive ligation reaction for the synthesis of large peptide dendrimers.

Peptide syntheses were performed on a CSBIO-536 automated synthesizer using CSBIO software. Purification of peptides was performed on a Waters instrument equipped with Vydac C18 reverse phase columns (size 25 x 2.2 cm i.d. and 25 x 1 cm i.d.). Analytical HPLC was performed on Shimadzu instruments including SCL-10A system controller, two LC-10AS pumps, SIL-10A auto injector, SPC-10A UV-VIS detector, and CR501 integrator. The analyses were carried out on a Vydac C18 reverse phase column (25 x 0.46 cm i.d.) at 1.0 ml/min monitoring at 225 nm. Eluents used were: A 0.046% TFA in water and B 0.039% TFA in 60% acetonitrile. The gradients used in the analyses are listed as follows: 1. oxime ligation: 0-1 min, 38% B, 1-21 min, linear gradient from 38-60%; 2. hydrazone ligation: 0-1 min, 38% B, 1-21 min, linear gradient from 38-64%;3. thiazolidine ligation: 0-1 min, 35% B, 1-21 min, linear gradient from 35-60% For amino acid analysis a fluorescence monitor was used for detection. The OPA/2-mercaptoethanol method was applied for amino acid analysis.[17] The molecular weight of the peptides was determined on a Kratos MALDI-MS III instrument.

Peptide VA20 with the sequence of VMEYKARRKRAAIHVMLALA was synthesized on the machine using the Fmoc/tBu strategy. p-Benzyloxybenzyl alcohol resin (Wang-resin) was used for synthesis. Coupling was accomplished by DCC/HOBT method with 2.5 equivalents of amino acids, and the Fmoc group was deprotected by 20% piperidine in DMF. The protected weak bases for ligation were introduced onto peptide through coupling 3 equivalents of Boc-$NHOCH_2COOH$, BOC-$NHNHCOCH_2CH_2COOH$, or Fmoc-Cys(Trt)-OH with BOP reagent. Final cleavage of peptides from the resin was performed with 90%TFA/6%thioanisole/3%ethanedithiol/1% anisole (50 mL/g resin) for 3 h. The resin was removed by filtration and the filtrates were concentrated *in vacuo*. After peptide products were precipitated with dry ether, they were filtered and further washed with dry ether. The precipitates were taken up in 100 mL of 10% acetic acid. Insoluble residues were removed by centrifugation. After lyophilization, these crude products were purified by RP-HPLC. The purified peptides were characterized by MALDI-MS and amino acid analysis.

MALDI-MS: NH_2OCH_2CO-VA20: 2402 ± 2.4 (Calcd. 2402); $NH_2NH(CH_2)_2CO$-VA20: 2444 ± 2.4 (Calcd. 2443); Cys-VA20: 2431 ± 2.4 (Calcd. 2432)

Amino acid analysis gave expected data for all peptides.

Boc-Lys(Boc)-Ala-OCH_3 (1). 1.40 g (10 mmol) of H-Ala-OCH_3HCl were suspended in a solution of 3.46 g (10 mmol) of Boc-Lys(Boc)-OH in 12 mL of DMF and 6 mL of DCM. The mixture was cooled at 0 °C and 4,42 g (10 mmol) of BOP and 2.84 g of DIEA (22 mmol) were added. The reaction mixture was stirred at 0 °C for 30 min and then at room temperature for 20 h. DCM and DMF were then removed *in vacuo*, the residue was taken up in 100 mL of ethyl acetate and the ethyl acetate phase was washed with saturated NaCl (2 x 15 mL), 2% $NaHSO_4$ (2 x 15 mL), 5% $NaHCO_3$ (3 x 15 mL) and water (3 x 15 mL). The organic phase was dried over anhydrous Na_2SO_4 and then concentrated to dryness. After recrystallization from ethyl acetate/hexanes, totally 4.37 g of dipeptide (1) were obtained (yield: 91.6%). MALDI-MS: 480.7 ± 0.4 (Calcd. for $M+H^+$ 480.6) elemental analysis: $C_{24}H_{37}N_3O_7$ (479.57) Found: C 60.31, H 7.92, N 8.48; Calcd: C 60.11, H 7.78, N 8.76

Boc-Lys(Boc)-Lys[Boc-Lys(Boc)]-Ala-OCH_3 (2). 0.96 g (2 mmol) of 1 were dissolved in 20 mL of 50% TFA/DCM. After stirring at room temperature for 20 min TFA and DCM were removed *in vacuo*. The residue was washed with dry ether (4 x 10 mL) and then dissolved in 20 mL of DMF. After the addition of 1.55 g (12 mmol) of DIEA, 1.44 g (4.2 mmol) of Boc-Lys(Boc)-OH, the mixture was cooled at 0 °C, and 1.86 g (4.2 mmol) of BOP were added to this solution in 0.5 min. The mixture was stirred at 0 °C for 30 min, then at room temperature for 24 h. The peptide was worked up as described in the preparation of dipeptide 1 preparation and final 1.65 g of title product were obtained (yield: 92.6%).

MALDI-MS: 910.4 ± 0.9 (Calcd. for $M+Na^+$ 911.1) elemental analysis: $C_{42}H_{77}N_7O_{13}$ (888.11) Found: C 57.01, H 8.88, N 10.68; Calcd: C 56.80, H 8.74, N 11.04

$\{[(CH_3O)_2CHCO]_2$-Lys$\}_2$-Lys-Ala-OCH_3 (3).

To the solution of 671 mg (5 mmol) of $(CH_3O)_2CHCOCH_3$ in 5 mL of methanol 10.5 mL of 0.5 N NaOH were added. The reaction was completed at in 2 hr by TLC monitoring. Methanol was removed *in vacuo* and the remaining aqueous solution was diluted to 25 mL with water. The solution was extracted with ethyl acetate (3 x 8 mL). The aqueous solution was concentrated to 10 mL and then lyophilized to dryness.

0.89 g (1 mmol) of 2 were dissolved in 20 mL of 50% TFA/DCM and the solution was stirred at room temperature for 20 min. After the removal of TFA and DCM, the residue was washed with dry ether (3 x 10 mL) and then dissolved in 10 mL of DMF. To this solution the powder obtained in A (5 mmol), 1.86 g (4.2 mmol) of BOP and 0.65 g (5 mmol) of DIEA were added. After stirring the reaction mixture for 20 h at room temperature, DMF was removed *in vacuo*. To the residue 10 mL of ethyl acetate were added and the solution was allowed to stand at 4 °C overnight. The precipitate was collected by filtration. The mother liquor was concentrated to dryness and then dissolved in 3 mL of water. After the addition of 4 drops of acetic acid, the precipitate (HOBt) was filtered off. The filtrate was concentrated to dryness. After washing the residue with dry ether, white powder was obtained and which was combined with the precipitate obtained from ethyl acetate above. The combined product was purified on a silica gel (40 g, 130-270 mesh, 60 A, Aldrich) column using $CHCl_3$:EtOAC:MeOH=60:25:15 as eluent. 680 mg of product were obtained.

Yield: 75.9%.

MALDI-MS: 919.2 ± 0.9 (Calcd. for M+Na^+ 918.99)

elemental analysis: $C_{38}H_{69}N_7O_{17}$ (896.00) Found: C 50.82, H 8.10, N 10.77; Calcd: C 50.94, H 7.76, N 10.94

CHOCO-Lys(CHOCO)-Lys(CHOCO-Lys(CHOCO))-Ala-OH **(4)** was synthesized as follows. To the solution of 134.4 mg (0.15 mmol) of peptide **3** in 10 mL water and 2 mL methanol, 1.8 mL of 0.1 N NaOH were added. After stirring at room temperature for 2 h, the hydrolysis reaction was completed according to TLC analysis. Methanol was removed in vacuo and the pH of solution was brought to 7.5 by adding 0.25 mL of 0.1 N HCl. After lyophilization, the white powder was dissolved in 1.5 mL of water. 500 mL (50 μmol) of this solution were taken out and to this solution 5 mL of conc. HCl were added. After stirring at room temperature for 3 min, the solution was concentrated to dryness in vacuo on a water bath of 35 °C. The residue was purified by RP-HPLC. After lyophilization 22.5 mg of pure aldehyde-MAP core were obtained. Yield 64.5%.

MALDI-MS: 699.0 (Calcd. 698.7)

The synthesis of MAPs through oxime, hydrazone, and thiazolidine ligation was performed as follows. All peptides were dissolved in water to give a 5 mM stock solution. The aldehyde-MAP core was dissolved in water as 5 mM stock solution.

For different ligation experiments in aqueous media, 50 μL of peptide stock solution were mixed with 50 μL of water, 100 μL of 0.2 M Na/HOAc buffer, 5 μL of aldehyde-MAP core solution. The final concentration was 1.25 mM for peptide and 0.125 mM for aldehyde-MAP core.

For ligation in 50% organic cosolvent, 50 μL of peptide stock solution were mixed with 50 μL 0.4 M Na/HOAc buffer, 100 μL of individual organic solvent , 5 μL of aldehyde-MAP core solution. The final concentration was 1.25 mM for peptide and 0.125 mM for aldehyde MAP core.

For ligation using 5 equivalents of peptide, 100 μL of peptide stock solution were mixed with 100 μL 0.2 M Na/HOAc buffer, 5 μL of aldehyde-MAP core solution. The final concentration was 2.5 mM for peptide and 0.125 mM for

aldehyde-MAP core.

All the reactions were followed by RP-HPLC analysis: 5 μL of reaction solution were taken at various time intervals and analyzed through RP-HPLC. The calculated rates of production formation are summarized in Fig. 14.

The ligation products were collected after separation by HPLC. Amino acid analysis of the products gave satisfactory results. MALDI-MS of final ligation products: VA20-Hdz-MAP: 10233 $\pm$ 10 (Calcd. 10230); VA20-Hdz-MAP: 10393 $\pm$ 10 (Calcd. 10395); VA20-Thz-MAP: 10347 $\pm$ 10 (Calcd. 10351). The amino acid analyses of the ligation products also gave the expected compositions.

Example 10.

The general approach is to exploit the selective reaction between a weak base and an aldehyde to form a stable conjugated phenyl hydrazone or ring compounds (Fig. 16). In this reaction, a weak base is distinguished from strong bases such as side chain and α-amines or the guanidino group of Arg in totally unprotected peptide fragments to serve as the only nucleophile reactive towards the aldehyde function under acidic conditions. The aldehyde group as an electrophile is distinguished from other functional groups in the peptide and is the only reactive moiety with the weak base in acidic pH. In this example, a weak base such as the 1,2-amino thiol of cysteine or 4-hydrazino benzoyl was placed at the N-terminal of a peptide fragment which was synthesized bv stepwise solid phase method. The peptide with the N-terminal weak base was deprotected from the resin with all side chain protecting groups removed. The purified unprotected peptide was then reacted in solution with the unprotected MAP core matrix containing multiple copies of aldehyde functions (a-oxoacyl), generated from the periodate oxidation of N-terminal serine of the $(Ser)_n$-Lysyl-ß-Ala scaffolding.

The core matrix containing either two or three levels of branched lysines and four or eight amino groups was synthesized by stepwise solid phase method using Fmoc-chemistry on a Wang resin (Fig. 17). Both $(Ser)_4$-$(Lys)_2$-Lys-ßAla (Ser_4-MAP) and $(Ser)_8$-$(Lys)_4$-$(Lys)_2$-Lys-ßAla (Ser_8-MAP) were obtained by TFA

cleavage from the peptide resin. Because only three and four coupling steps, respectively, were involved in these syntheses, both products were obtained in high purity (>95%) and were used without further purification. The same compounds could also be achieved by Boc-chemistry. To functionalize the MAP core matrix with reactive aldehyde groups, the lysinyl-resin was capped with Fmoc-Ser(t-Bu) to give an 2-amino alcohol moiety which was transformed by periodate oxidation to generate an α-oxoacyl group. Oxidation of Ser-MAP to CHO-MAP containing four or eight copies of α-oxoacyl groups was achieved using two-molar excess of sodium meta-periodate in aqueous buffer at pH 7 for 5 min. The reaction was quenched with 4 fold excess of ethylene glycol. No side products were observed by C18 RP-HPLC. Because formaldehyde was generated during the quenching as well as from the reaction, the (CHO)-MAP forms were purified by C18 RP-HPLC. Both forms of Ser-MAP and CHO-MAP were characterized by laser desorption mass spectrometric (LD-MS) analysis to give the expected MW. The aldehydic MAP cores were used immediately for the conjugation reaction (Fig. 18).

Two unprotected peptide fragments were selected for the conjugation reactions. The first peptide fragment, CA-16 (CNYNKRKRIHIPGPRA-NH_2), contains 16 residues and is rich in basic amino acids. This peptide is the neutralizing epitope and part of the third variable region (V3 loop) of gp 120 of human immunodeficient virus, MN strain. Rusche, J.R., Javaherian, K., McDanal, Pedro, J., Lynn, D.L., Grimaila, R., Langlois, A., Gallo, R.C., Arthur, L.O., Fischinger, P.J., Bolognesi, D.P., Putney, S.D. and Matthews, T.J. (1988) *Proc. Natl. Acad. Aci. U.S.A.* **85**, 3198-3202; Goudsmit, J., Debouck, C., Meloen, R.H., Smith, L., Bakker, M., Asher, D.M., Wolff, A.F., Gibbs Jr, C.J. and Gajdusek, D.C. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4478-4482; Javaherian, K., Langlois, A.J., McDanal, C., Ross, K.L., Eckler, L.I., Jellis, C.L., Profy, A.T., Rusche, J.R., Bolognesi, D.P., Putney, S.D. and Matthews, T.J. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6768-72; Devash, Y., Calvelli, T.A., Wood, D.G., Reagan, K.J. and Rubinstein, A. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 3445-49; Takahashi, H., Cohen, J., Hosmalin, A., Cease, K.B., Houghten, R., Cornette, J.L., DeLisi, C., Moss, B., Germain, R.N. and Berzofsky, J.A. (1988)

Proc. Natl. Acad. Sci. U.S.A. 85, 3105-3109. A cysteine was placed at the N-terminal and was ideally suited for the ligation to the aldehyde groups of the core matrix in the formation of a 5-member thiazolidine ring. The second peptide fragment, SR-10 (SSQFQIHGPR) contains 10 residues and is an autoimmune epitope from the ZP3 glycoprotein. Rusche, J.R., Javaherian, K., McDanal, Pedro, J., Lynn, D.L., Grimaila, R., Langlois, A., Gallo, R.C., Arthur, L.O., Fischinger, P.J., Bolognesi, D.P., Putney, S.D. and Matthews, T.J. (1988) *Proc. Natl. Acad. Aci. U.S.A.* 85, 3198-3202; Goudsmit, J., Debouck, C., Meloen, R.H., Smith, L., Bakker, M., Asher, D.M., Wolff, A.F., Gibbs Jr, C.J. and Gajdusek, D.C. (1988) *Proc. Natl. Acad. Sci. U.S.A.* 85, 4478-4482; Javaherian, K., Langlois, A.J., McDanal, C., Ross, K.L., Eckler, L.I., Jellis, C.L., Profy, A.T., Rusche, J.R., Bolognesi, D.P., Putney, S.D. and Matthews, T.J. (1989) *Proc. Natl. Acad. Sci. U.S.A.* 86, 6768-72; Devash, Y., Calvelli, T.A., Wood, D.G., Reagan, K.J. and Rubinstein, A. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87, 3445-49; Takahashi, H., Cohen, J., Hosmalin, A., Cease, K.B., Houghten, R., Cornette, J.L., DeLisi, C., Moss, B., Germain, R.N. and Berzofsky, J.A. (1988) *Proc. Natl. Acad. Sci. U.S.A.* 85, 3105-3109. It was prepared with a phenyl hydrazine group at the amino terminus so it could form hydrazone linkages with the CHO-MAP core. Both fragments, CA-16 and hydrazinobenzoyl(Hob)-SR-10 were prepared without difficulty by solid-phase method using Boc-chemistry and purified bv C18 RP-HPLC to homogeneous products (Fig. 18 and 19). The phenyl hydrazine group was introduced to SR-10 through an amide linkage to the amino group of the peptide with 4-Boc-hydrazinobenzoic acid using DCC/HOBt at the final step of the synthesis. The Hob group was found to be stable to the usual cleavage conditions of using TFA (Fmoc) chemistry and HF (Bio chemistry).

All solvents used for conjugation reactions were deaerated and purged with argon. DMF, DCM (both biotechnology.grade), acetonitrile and MeOH (HPLC-grade) were obtained from Fisher Scientific; DCC, p-cresol and piperidine (Aldrich Chem.); and TFA (Halocarbon). Boc- and Fmoc-amino acids were purchased from Bachem (Torrance, California). All other chemicals were the purest grade available. All peptides and conjugate products were analyzed and purified on

Shimadzu and Waters HPLC'S. Mass spectrometric analysis was determined by matrix assisted laser-desorption mass-spectrometry (Kratos).

CA-16 carboxamide (CNTNKRKRIHIPGPRA-NH_2) was synthesized by the solid-phase using the Boc/Bzl strategy. The synthesis was started from the 4-methybenzhydrylamine-resin (0.54 mmol/g). Side chain protecting group were: Arg(Tos), Cys(4-MeBzl), His(Dnp), Lys(ClZ), Ser(Bzl), Thr(Bzl) and Tyr(BrZ). All amino acids were protected with Boc on the Nα-terminus. All couplings were performed in DCC except for Boc-Arg and Boc-Asn, which were carried out in HBTU and DIEA. After synthesis, the protected peptide resin was treated with 10% thiophenol in DMF (v/v) for totally 24 hr to remove the N^{in}-Dnp protecting group on His. The dried peptide-resin was deprotected and cleaved by High HF (p-cresol) for 1.5 hr at O°C. After cleavage, the crude peptide was purified by preparative C18-RP HPLC. The purified peptide was analyzed by matrix assisted laser-desorption mass-spectrometry (LDMS); [CA-16] (cal/found) (1884/1884).

Solid phase peptide synthesis of SR-10 (SSQFQIHGPR) was performed on a Pam resin (0.41 mmol/g). The procedure was the same as described above. At the completion of the synthesis, the resin was coupled with 4 molar excess of 4-Boc-NHNH-C_6H_4-CO_2H via DCC/HOBt in DMF/DCM (1:1, v/v). LDMS analysis gave calcd/found (1290.5/1291).

Ser_4-Lys_2-Lys-ßAla and Ser_8-Lys_4-Lys_2-Lys-ßAla were synthesized by a stepwise solid phase method using the Fmoc/t-Bu strategy on a Wang-resin (0.5 mmol/g). Attachment of Fmoc-ßAla (6 equiv) to the resin was performed in DCC (3 equiv) and 10% DMAP (0.1 mmol) in DMF/DCM (1:1, v/v). Synthesis of 2 and 3 levels of the lysines were achieved using 3 molar excess of Fmoc-Lys(Fmoc) and DCC in DMF/DCM (1:1, v/v). After Fmoc deprotection, Fmoc-Ser(t-Bu) was coupled to the lysine containing MAP core-matrix via DCC in DMF/DCM. Both Ser_4-Lys_2-Lys-ßAla and Ser_8-Lys_4-Lys_2-Lys-ßAla were cleaved from the resin by TFA and lyophilized for immediate use without purification. Conversion of the Ser-MAP to glyoxyl-MAP was achieved by oxidation with meta-periodate (46.8 μmol, 10 mg) and Ser_n-MAP (n=4, 6.1 μmol, 5 mg,; n=8, 2.97 μmol, 5 mg) in 0.01 M sodium phosphate-buffer, pH 7 (n=4, 300 μl; n=8, 210 μl). After the

reaction was mixed for 5 min at room temperature, it was quenched by adding ethylene glycol (93.5 μmol, 5.3 μl) to remove the excess of sodium meta-periodate. The mixture was purified by C18-RP HPLC with a 30-min linear gradient of 0-60 % buffer B. Buffer A contained 100 % water, 0.045 % TFA and Buffer B contained 60 % CH_3CN, 0.039 % TFA. Both MAP core matrixes were used immediately for the conjugation reaction after HPLC purification. LDMS analysis gave the expected molecular weights, (Ser_4-MAP, calcd/found (821/822), $(HCO\text{-}CO)_4$-MAP, calcd/found (697.7/698), Ser_8-MAP, calcd/found (1682.5/1683), $(HCO\text{-}CO)_8$-MAP, calcd/found (1434.5/1436).

The unprotected peptide fragment CA-16 and (CHO)4-MAP were assembled to form a tetrameric branched peptide in aqueous buffer at pH 5 using a 4 molar equivalents of peptide for each oxoacyl site. To investigate the optimal condition for the formation of the 5-member thiazolidine ring between α-oxoacyl group at the MAP core and N-terminal cysteine, the effects of organic co-solvent, temperature, and antioxidants such as EDTA (ethylenediaminetetraacetic acid) to prevent disulfide formation were examined (Figure 20).

Since relatively large excess of monomers were used, the pseudo first order rate was evaluated based on the formation of 50% of the conjugation product monitored by C8 RP-HPLC. In these experiments, all samples were deaerated and saturated with argon to exclude oxygen in the solution. However, it was found that this condition was insufficient to prevent disulfide oxidation of the peptide CA-16 that consumed the starting material and lowered the yield. The use of an anti-oxidant or chelating agent such as EDTA, which removed metal cations in the aqueous solution that catalyzed disulfide oxidation, was studied. EDTA (0.8 mM) was sufficient to reduce disulfide formation of CA16 to < 5% and increased the $t_{½}$ of the rate formation from 15 to 10.7 hr. Thus, EDTA was used in all subsequent experiments to minimize disulfide formation.

The rate of thiazolidine ring formation was dependent on three factors. The most significant appeared to be the organic co-solvent which gave a rate increase of about 3 fold when 60% of DMF was present at 22°C. As expected, increased temperature from 22°C to 37°C enhanced the rate about 2 fold in H_2O or 60 % DMF, and nearly

5 and 3 fold at 50°C, respectively. Finally, the increase in concentration from 4.8 M to 8 M had slight beneficial effects. The combination of organic co-solvent and elevated temperature provided not only rate acceleration but also reduced the danger of aggregation that leads to precipitation.

Peptide CA- 16 was also conjugated to the CHO-MAP core containing eight aldehyde groups to give a highly compact and dense octameric peptide dendrimer. This conjugation reaction was performed in aqueous buffer containing EDTA at 37°C and was completed within 12 hrs with only one major peak (Fig. 18) and a side product < 5% from the oxidation of CA-16 to CA-16 disulfide dimer was observed. The products obtained from conjugation of CA-16 to $(CHO)_4$-MAP and $(CHO)_8$-MAP to give the tetrameric and octameric branched dendrimers were analyzed and gave the correct MW and amino acid analysis (Figure 21). Furthermore, despite their high MWs, they gave a single peak in RP-HPLC. These results show that the reaction is highly specific and that no evidence of side reaction between the α-oxoacyl group with the side chain nucleophiles can be found.

Conjugated phenyl hydrazine has a pKa of about 4.5 and forms a stable conjugated hydrazone with an alkyl aldehyde. Thus, in this approach, the unprotected peptide Hob-SR-10 was assembled on the $(CHO)_4$-MAP core matrix at pH 5 to form the tetrameric peptide dendrimer. The hydrazone formation in aqueous solution between Hob-SR-10 and $(CHO)_4$-MAP was rapid and selective and dependent on the stoichimetric ratio of peptide to each aldehyde group. With an equal molar ratio of peptide and α-oxoacyl group, the reaction was completed in 5 hr but accelerated to 1 hr and 10 min, respectively, when the stoichimetric ratio of peptide and aldehyde was increased to 1.5 and 2. The progress of the reaction was conveniently monitored by analytical RP-HPLC. Because phenyl hydrazine and the conjugation product, phenyl hydrazone, differ in their absorption maximum by 70 nm, it was monitored by uv spectrometry (Fig. 22). To achieve eight copies of SR1O on the MAP core matrix, the reaction was performed in DMF:H20 (1:1, v/v) to prevent precipitation of the intermediates to give the octameric SR-10 MAP. This reaction was monitored by RP-HPLC and only one single product was formed after 12 hr (Fig. 19). The molecular mass for $(SR\text{-}10)_4$-Hab-MAP and $(SR\text{-}10)_8$-Hab-MAP, respectively, was

confirmed by LD-MS (Fig. 21) in addition to amino acid analysis.

Peptide CA-16 (18 μmol, 34 mg) was dissolved in 0.02M NaOAc buffer containing 0.008M EDTA, pH 5 (2.8 mi) and the (HCOCO)4-MAP solution (1.13 μmol) collected from the HPLC was added. The solution was adjusted to pH 5 with pyridine. The deaerated solution was kept under argon and in dark for 4 hrs at 50°C. The conjugation reaction was monitored by analytical C8-RP HPLC using a linear gradient of solvent B, 2 %/min, starting from 0 % and with a flow of 1 ml/min. Buffer A contained 5 % CH_3CH, 0.045 % TFA and Buffer B contained 60 % CH_3CN, 0.039 % TFA. The conjugated product was purified by semi-preparative HPLC (C18-RP coumn) under isocratic conditions at a flow of 2 ml/min. The yield of $(peptide)_4$-thiazolidine-MAP was 9.2 mg. The purified peptide-MAP conjugate was analyzed by LDMS, ($(NA\text{-}15)_4$-Thz-MAP) calcd/found (8160.5/8162).

Reaction between $(HCOCO)_8$-MAP and CA-16 was achieved using the conditions described above, except that the deaerated solution was kept under argon and in dark overnight at 37°C. The progress of the reaction was monitored by analytical C8-RP HPLC using a linear gradient of solvent B, 0.94 %/min, starting from 20 % and with a flow of 1 ml/min. Buffer A contained 0 % CH_3CH, 0.045 % TFA and Buffer B contained 60 % CH_3CN, 0.039 % TFA. LDMS analysis gave calcd/found (16362.5/16363).

The $(HCO\text{-}CO)_4$-MAP solution (1.65 μmol) collected from the HPLC was mixed with 4-hydrazino-benzoyl-SR10 (13.2 μmol, 17 mg) and the solution was adjusted to pH 5 with 0.02 M sodium acetate buffer, pH 5. The deaerated solution was kept under argon and in dark for 1 hr at room temperature. The reaction was monitored by analytical C8-RP HPLC using a linear gradient of solvent B, 2 %/min, starting from 0 % and with a flow of 1 ml/min. The conjugated product was purified by C18-RP HPLC using isocratic conditions. The yield of the $(SR\text{-}10)_4$-Hyz-MAP was 6 mg. The peptide-MAP conjugates was analyzed by LDMS, calcd/found (5788/5789.1).

For the conjugation reaction between 4-hydrazinobenzoyl-SR-10 and $(HCOCO)_8$-MAP, H_2O/DMF (1:1), pH 5 was used in 0.02 M sodium acetate buffer. Three mol equivalents of peptide was used for each a-oxoacyl group on $(CHO)_8$-MAP. The reaction was carried out at

37°C for 12 hr. LDMS analysis gave calcd/found (11614.5/11615).

The stability of the 5-member thiazolidine rings and phenyl hydrazone bonds at pH 5-8 were investigated under a wide range of pH in aqueous conditions. The peptide dendrimer derived from ligating CA-16 with the α-oxoacyl MAP core matrix, $(NA\text{-}15)_4$-Thz-MAP, was incubated at 37°C at pH 5, 6, 7 and 8 and samples were withdrawn at 12 hr intervals for 5 days and analyzed by HPLC. The ligation site of thiazolidine ring in $(NA\text{-}15_4$-Thz-MAP showed no evidence of significant hydrolysis after 5 days at these pH values and remained as a single symmetrical peak in RP-HPLC (Fig. 18). Similarly when $(SR\text{-}10)_4$-Hab-MAP was incubated at 37°C at pH 6, 7, 7.4 and 8 and monitored by HPLC over 2 days, the phenyl hydrazone linkage in $(SR\text{-}10)_4$-Hab-MAP remained stable (Fig. 19). Alkyl hydrazone linkages are usually susceptible to hydrolysis and required stabilization by reduction with sodium cyanoborohydride. However, due to the aromatic character in the phenyl hydrazone linkage, it is more resistant to hydrolysis than the alkyl hydrazone and does not require the extra step of reduction for stabilization. These results show that peptide dendrimers with these linkages have substantial stability in aqueous media to be useful at the physiological pHs as drug carriers, vaccines and diagnostic reagents.

Examples 11 to 13 illustrate the application of domain ligation strategy in the site-specific modification of proteins.

Conceptually, site-specific modification of proteins is similar in both principle and practice to the ligation of peptides because proteins are used with many unprotected side chains and the site-specificity of the domain ligation operates. In our examples, the proteins are obtained from the recombinant methods but proteins purified from the natural sources can also be used. Ideally, a protein with an amino terminal cysteine will be useful as a weak base and a protein with an amino terminal serine or threonine can serve as an aldehyde, which can be converted to an aldehyde function by sodium metaperiodate oxidation. This oxidation is relatively selective at pH 7 for the 1,2-aminoethanol such as serine and threonine and is >300 fold faster than the oxidation of 1,2-diol found in the carbohydrates. With the recombinant technology, any proteins can be engineered to contain one of these residues. For naturally isolated proteins lacking in one of these residues, the 1,2- or 1,3-amino thiol moiety can be introduced

via a substituted thiolactone (Fig. 23). Unsubstituted thiolactone is too unhindered and gives very poor yield because of self-condensation. The thiolactone forms reversible Schiff base with the aldehyde component that gives nonproductive reaction; but, when the thiolactone reacts with the N-terminal of the protein (some minor reaction with the side chain amine of lysine) the thiol moiety is then liberated to give the ring compound similar to the ring formation of the domain ligation strategy and a stable product.

Example 11 illustrates the principle of domain ligation to modify gp120, which is the surface protein of HIV-1 and a principle target of vaccine development. The site-specific modification by a lipid module such as tripalmitoyl glceryl cysteine (P3C)-Lys(Cys) or Cys-Lys(Pal)-D-Lys(Pal) can be achieved by oxidation of the amino terminal threonine of gp120 to an aldehyde.

Example 12 and 13 illustrate the same principle by modifying any protein to contain an 1,2- or 1,3-amino thiol by reacting with an N-substituted thiolactone in the presence of an aldehyde.

Example 11

The site-specific lipidation of gp120 using a lipid module and oxidation of the N-terminal threonine of gp120 to form an aldehyde was performed.

Recombinant derived gp120 of HIV-1, IIIB strain (1 mg) in 0.1 M sodium phosphate buffer at Ph 7 (2 Ml) was treated with 2 molar equivalents of sodium metaperiodate to oxidize the N-terminal threonine to an glyoxyl aldehyde. The reaction was stopped after 5 min and quickly dialyzed in the same buffer (1 liter) to remove the oxidant and formaldehyde generated. The oxidized gp 120 was then reacted with a biphasic solution (2 ml of methylene chloride-dimethylformaide at 1:1 volume ratio) for 24 hr containing 1.5 equivalent of the following a lipid module containing a cysteine attached to the side chain of an lysinyl peptide: tripalmitoyl glyceryl Cys-Lys-Cys or Cys-Lys(Pal)-D-Lys(Pal)-OH. At the completion of the reaction, the organic phase was evaporated by reduced pressure and the resulting aqueous phase was subjected to purification by high performance gel permeation chromatography to obtain the lipidated gp120.

Example 12

The site-specific lipidation of gp120 using a lipid module and a secondary amine thiolactone was performed.

Recombinant derived gp120 (1 mg) in 0.1 M sodium phosphate buffer at Ph 7 (2 Ml) was treated with 2 equivalents a substituted thiolactone and a lipid module in a biphasic solution (see Example 11) for 24 hr. At the completion of the reaction the organic phase was evaporated by reduced pressure and the lipidated gp120 was purified by high performance gel permeation chormatography.

Example 13

The site-specific pegylation of Interleukin-2 by MeO-PEG-CHO and a secondary amine thiolactone was performed.

The reaction condition and work-up was similar to Example 12.

Thus, although there have been described particular embodiments and examples of the present invention, it is not intended that such references be construed as limitations upon the scope of this invention except as set forth in the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Tam, James P.

(ii) TITLE OF INVENTION: Domain Ligation Strategy to Engineer Proteins with Unusual Architectures

(iii) NUMBER OF SEQUENCES: 2

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Waddey & Patterson

(B) STREET: L & C Tower, 27th Floor, 401 Church St.

(C) CITY: Nashville

(D) STATE: Tennessee

(E) COUNTRY: U.S.A.

(F) ZIP: 37219

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette, 3.50 inch, 800 kB storage

(B) COMPUTER: IBM PC/XT/AT compatible

(C) OPERATING SYSTEM: MS-DOS (version 5.0)

(D) SOFTWARE: WordPerfect 5.1/WordPerfect Editor

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 08/081,412

(B) FILING DATE: 21-JUN-1993

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Waddey, I.C., Jr.

(B) REGISTRATION NUMBER: 25,180

(C) REFERENCE/DOCKET NUMBER: 0216-9334

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (615) 242-2400

(B) TELEFAX: (615) 242-2221

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 15 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: yes

(v) FRAGMENT TYPE: entire synthesized peptide

(ix) FEATURE:

(A) NAME/KEY: a biologically non-significant peptide synthesized for use in demonstration of the domain ligation strategy's effectiveness in the synthesis of proteins

(D) OTHER INFORMATION: a synthetic peptide without specific biological localization or significance

(x) PUBLICATION INFORMATION:

(A) AUTHOR: Tam, James P

(B) TITLE: A Chemical Ligation Approach to the Synthesis of Proteins Using Unprotected Peptide Segments. Concept and Model Study.

(C) JOURNAL: J. Amer. Chem. Soc. (1993) In press.
(K) RELEVANT RESIDUES In SEQ ID NO:1: FROM 9 TO 11
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Cys Tyr Thr Ser Gly Cys Val Arg Ala Pro Thr Phe Asp Leu Lys
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:2:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal fragment
(ix) FEATURE:
(A) NAME/KEY: Thr-33 TGF alpha
(C) IDENTIFICATION METHOD: Synthesis of the TGF alpha analog in vitro by means of the domain ligation strategy
(D) OTHER INFORMATION: a synthetic analog of TGF alpha produced by means of the domain ligation strategy
(K) RELEVANT RESIDUES: In SEQ ID NO:2: FROM 32 TO 34
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Val Val Ser His Phe Asn Lys Cys Pro Asp Ser His Thr Gln Tyr Cys
1 5 10 15

Phe His Gly Thr Cys Arg Phe Leu Val Gln Glu Glu Lys Pro Ala Cys
20 25 30

Thr Cys His Ser Gly Tyr Val Gly Val Arg Cys Glu His Ala Asp Leu
35 40 45

Leu Ala
50

(2) INFORMATION FOR SEQ ID NO: 3:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: yes
(v) FRAGMENT TYPE: N-terminal fragment

(ix) FEATURE:
(A) NAME/KEY: a biologically non-significant peptide synthesized for use in demonstration of the domain ligation strategy's effectiveness in the synthesis of proteins
(D) OTHER INFORMATION: a synthetic peptide without specific biological localization or significance
(x) PUBLICATION INFORMATION:
(A) AUTHOR: Tam, James P
(B) TITLE: A Chemical Ligation Approach to the Synthesis of Proteins Using Unprotected Peptide Segments. Concept and Model Study.
(C) JOURNAL: J. Amer. Chem. Soc. (1993) In press.
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Cys Tyr Thr Ser Gly Cys Val Arg
1 5

(2) INFORMATION FOR SEQ ID NO: 4:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: yes
(v) FRAGMENT TYPE: N-terminal fragment
(ix) FEATURE:
(A) NAME/KEY: a biologically non-significant peptide synthesized for use in demonstration of the domain ligation strategy's effectiveness in the synthesis of proteins
(D) OTHER INFORMATION: a synthetic peptide without specific biological localization or significance
(x) PUBLICATION INFORMATION:
(A) AUTHOR: Tam, James P
(B) TITLE: A Chemical Ligation Approach to the Synthesis of Proteins Using Unprotected Peptide Segments. Concept and Model Study.
(C) JOURNAL: J. Amer. Chem. Soc. (1993) In press.
(K) RELEVANT RESIDUE: In SEQ ID NO:4: Residue No. 9
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Cys Tyr Thr Ser Gly Cys Val Arg Ala
1 5

(2) INFORMATION FOR SEQ ID NO: 5:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: yes
(v) FRAGMENT TYPE: C-terminal fragment
(ix) FEATURE:
(A) NAME/KEY: a biologically non-significant peptide synthesized for use in demonstration of the domain ligation strategy's effectiveness in the synthesis of proteins
(D) OTHER INFORMATION: a synthetic peptide without specific biological localization or significance
(x) PUBLICATION INFORMATION:
(A) AUTHOR: Tam, James P
(B) TITLE: A Chemical Ligation Approach to the Synthesis of Proteins Using Unprotected Peptide Segments. Concept and Model Study.
(C) JOURNAL: J. Amer. Chem. Soc. (1993) In press.
(K) RELEVANT RESIDUE: In SEQ ID NO:5: Residue No. 1
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Thr Phe Asp Leu Lys
1 5

(2) INFORMATION FOR SEQ ID NO:6:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(A) DESCRIPTION: a synthetic version of the N-terminal 32 amino acids of TGF alpha which are employed in the domain ligation strategy method of Thr-33 TGF alpha synthesis
(iii) HYPOTHETICAL: no
(v) FRAGMENT TYPE: N-terminal fragment
(ix) FEATURE:
(A) NAME/KEY: TGF alpha analog
(C) IDENTIFICATION METHOD: Synthesis of the TGF alpha analog in vitro by means of the domain ligation strategy
(D) OTHER INFORMATION: a synthetic analog of TGF alpha produced by means of the domain ligation strategy
(K) RELEVANT RESIDUES: In SEQ ID NO:6: Residue No. 32
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Val Val Ser His Phe Asn Lys Cys Pro Asp Ser His Thr Gln Tyr Cys
1 5 10 15

Phe His Gly Thr Cys Arg Phe Leu Val Gln Glu Glu Lys Pro Ala Cys
20 25 30

(2) INFORMATION FOR SEQ ID NO:7:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(A) DESCRIPTION: a synthetic version of the C-terminal 18 amino acids of TGF alpha which are employed in the domain ligation strategy method of Thr-33 TGF alpha synthesis
(iii) HYPOTHETICAL: no
(v) FRAGMENT TYPE: C-terminal fragment
(ix) FEATURE:
(A) NAME/KEY: TGF alpha analog
(C) IDENTIFICATION METHOD: Synthesis of the TGF alpha analog in vitro by means of the domain ligation strategy
(D) OTHER INFORMATION: a synthetic analog of TGF alpha produced by means of the domain ligation strategy
(K) RELEVANT RESIDUES: In SEQ ID NO:7: Residue No. 1
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Thr Cys His Ser Gly Tyr Val Gly Val Arg Cys Glu His Ala Asp Leu
1 5 10 15

Leu Ala

CLAIMS

What I claim is:

1. A method of chemically ligating a C-terminal carboxylic group of a first peptide segment and a N-terminal amino group of a second peptide segment, said method comprising:

a) introducing a masked aldehyde ester function to said carboxylic group of said first peptide segment;

b) converting said masked aldehyde ester function to its aldehyde on said carboxylic group of said first peptide segment;

c) reacting a weak base attached to said N-terminal amino acid of said second peptide segment with said aldehyde to form a bonding structure whereby said carboxylic group and said amino group are brought together to form an amide bond through O to N-acyl rearrangement.

2. The method according to claim 1, wherein said first peptide segment is physically separate and unattached from said second peptide segment.

3. The method described in claim 1, wherein said first and said second peptide segments comprise opposite ends of the same polypeptide chain.

4. The method according to claim 1, wherein said carboxylic group of said first peptide segment is esterified to an masked α-hydroxyaldehyde to form a masked ester α-aldehyde.

5. The method according to claim 1 where said weak base includes a 1,2-amino thiol of cysteine, a 1,2-amino ethanol of serine, a 1,2-amino ethanol of threonine, an aminooxyacetyl function, a mono-hydrazine succinyl function, or 4-hyrazino benzoyl.

6. The method according to claim 1, where said bonding structure includes a thiazolidine ring, an oxazolidine ring, an oxime, a hydrazone, or a phenylhydrazone.

7. The method according to claim 1, wherein said masked aldehyde ester is introduced to said carboxylic group of said first peptide segment by an enzyme-catalyzed coupling reaction.

8. The method according to claim 1, wherein said masked aldehyde ester function is introduced to said carboxylic group of said first peptide by a procedure selected from

the group comprising reverse proteolysis, solid phase synthesis, or lipase ligation.

9. The method according to claim 1, wherein said masked aldehyde ester has the formula:

$$-O-(CR_1R_2)n-CHO$$

wherein R_1 R_2 are hydrogen, simple alkyl or substituted aromatic functional groups and n=1 to 9.

10. The method according to claim 9, wherein said masked ester aldehyde comprises a simple alkyl aldehyde, an α-formyl methyl ester, an α-formyl ethyl ester or a ß,ß,ß-dimethyl formyl ethyl ester.

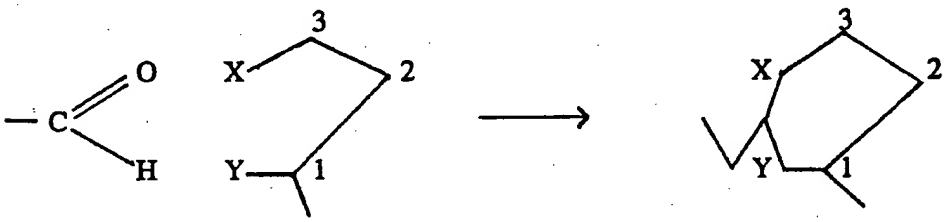
11. The method according to claim 1, wherein said method is practiced in a single reaction vessel.

12. The method according to claim 1, wherein said method relies solely on pH changes to initiate changes in the reaction.

13. The method according to claim 1, wherein said method involves a reaction which occurs entirely in an aqueous solution.

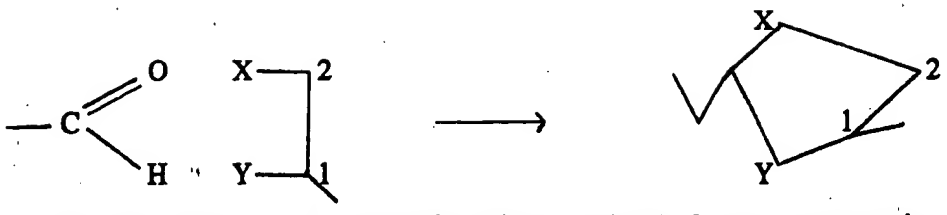
14. The method according to claim 1, wherein said method involves no intermediate purification steps.

15. The method according to Claim 1 wherein a carboxylic group is chemically ligated with an amino group by the interaction of a masked aldehyde ester with a 1,3-disubstituted α-amino acid which is able to form a 6-member ring having the general structure



where X= N, O, S or an aromatic amine and Y=N, O, S or an aromatic amine.

16. The method according to Claim 1 wherein a carboxylic group is chemically ligated with an amino group by the interaction of a masked aldehyde ester with a 1,2-disubstituted α-amino acid which is able to form a 5-member ring having the general structure



where X= N, O, S or an aromatic amine and Y=N, O, S or an aromatic amine.

17. A method for synthesizing a peptide dendrimer by chemically ligating an unprotected peptide to a scaffolding matrix, said method comprising:

(a) converting an alcohol moiety at N-terminal amino acid on a branch of said scaffolding matrix to an aldehyde;

(b) reacting a weak base attached to the N-terminal amino acid of said peptide with said aldehyde to form a bonding structure; and,

(c) allowing said reaction to proceed until all said branches of said scaffolding matrix are occupied by a molecule of said peptide.

18. The method described in claim 17 where said scaffolding matrix is octavalent lysinyl multiple antigen protein.

19. The method described in claim 17 where said scaffolding matrix is tetravalent lysinyl multiple antigen protein.

20. The method described in claim 17 where said N-terminal amino acid of each said branch of said scaffolding matrix comprises serine.

21. The method described in claim 17 where said aldehyde is an α-oxoacyl aldehyde.

22. The method described in claim 17 where said weak base includes a 1,2-amino thiol of cysteine, an aminooxyacetyl function, a mono-hydrazine succinyl function, or 4-hyrazino benzoyl.

23. The method described in claim 17 where said bonding structure comprises a thiazolidine ring, an oxime, a hydrazone, or a phenylhydrazone.

24. The method described in claim 17 where said peptide comprises a 24-amino acid fragment from V3-Loop of gp120 of HIV-1, MN strain.

25. The method described in claim 17 where said peptide comprises a 16-amino acid fragment from V3-Loop of gp120 of HIV-1, MN strain.

26. The method described in claim 17 where said peptide comprises a 20-amino

acid fragment derived from a surface protein of feline leukemia virus, VA-20.

27. The method described in claim 17 where said peptide comprises a 10-amino acid fragment from ZP3 glycoprotein.

28. A peptide dendrimer produced by the method described in claim 17.

29. The peptide dendrimer described in claim 28, where said scaffolding matrix is octavalent lysinyl multiple antigen protein.

30. The peptide dendrimer described in claim 28, where said scaffolding matrix is tetravalent lysinyl multiple antigen protein.

31. The peptide dendrimer described in claim 28, where said peptide comprises a 24-amino acid fragment from V3-Loop of gp120 of HIV-1, MN strain.

32. The peptide dendrimer described in claim 28, where said peptide comprises a 16-amino acid fragment from V3-Loop of gp120 of HIV-1, MN strain.

33. The peptide dendrimer described in claim 28, where said peptide comprises a 20-amino acid fragment derived from a surface protein of feline leukemia virus, VA-20.

34. The peptide dendrimer described in claim 28, where said peptide comprises a 10-amino acid fragment from ZP3 glycoprotein.

35. A method for site-specific modification of a protein, said protein engineered to have at its N-terminus an amino acid with an alcohol moiety, by chemically ligating said protein to a lipid module, said lipid module comprised of a lipid and a peptide, said method comprising:

(a) converting said alcohol moiety at said N-terminal amino acid of said protein to an aldehyde; and

(b) reacting a weak base attached to a side chain amino acid of said peptide of said lipid module with said aldehyde to form a ring.

36. The method described in claim 35, where said N-terminal amino acid of said protein is threonine.

37. The method described in claim 35, where said N-terminal amino acid of said protein is serine.

38. The method described in claim 35, where said aldehyde is a glyoxylyl aldehyde.

39. The method described in claim 35, where said lipid module is tripalmitoyl glyceryl cysteine.

40. The method described in claim 35, where said side chain amino acid of said protein is cysteine.

41. The method described in claim 35, where said weak base comprises 1,2-amino thiol of cysteine.

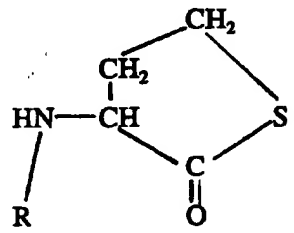
42. The method described in claim 35, where said protein is recombinant derived gp120 of HIV-1, IIIb strain.

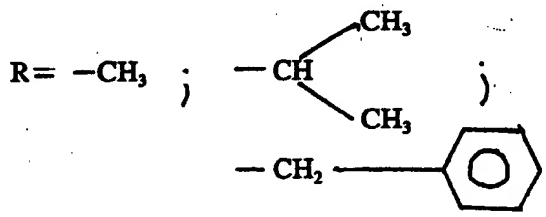
43. A method for site-specific modification of a protein comprising:

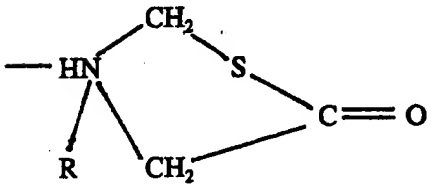
(a) introducing a 1,2- or 1,3-amino thiol ring moiety by reacting a secondary amine thiolactone with the N-terminal of said protein; and,

(b) reacting said thiol moiety with an aldehyde attached to a modifying molecule.

44. The method described in claim 43 where said secondary amine thiolactone comprises the following:







45. The method described in claim 43 where said modifying molecule is a lipid module.

46. The method described in claim 43 where said protein is recombinant gp 120.

47. The method described in claim 43 where said modifying molecule is MeO-PEG-CHO.

48. The method described in claim 43 where said protein is interleukin-2.

49. A protein modified by the method described in Claim 35.

50. A protein modified by the method described in Claim 43.

AMENDED CLAIMS
[received by the International Bureau on 17 November 1994 (17.11.94)]
original claims 1-50 replaced by amended claims 1-76 (9 pages)]

1. A method of chemically ligating a first macromolecule to a second molecule, said method comprising reacting a weak base moiety of the first macromolecule with an aldehyde moiety of the second molecule.

2. The method described in claim 1 wherein said reaction of the weak base moiety with the aldehyde moiety is performed under acidic conditions.

3. The method described in claim 1 wherein the first macromolecule is a protein, a DNA molecule, or a peptide scaffolding matrix.

4. The method described in claim 1 wherein the second molecule is a macromolecule.

5. The method according to claim 1, wherein the weak base moiety is selected from the group consisting of a 1,2-amino thiol of cysteine, a 1,2-amino ethanol of serine, a 1,2-amino ethanol of threonine, an aminooxyacetyl function, a mono-hydrazine succinyl function, or 4-hyrazino benzoyl.

6. A method of chemically ligating a first macromolecule to a second molecule, said method comprising reacting an aldehyde moiety of the first macromolecule with a weak base moiety of the second molecule.

7. The method described in claim 6 wherein said reaction of the weak base moiety with the aldehyde moiety is performed under acidic conditions.

8. The method described in claim 6 wherein the first macromolecule is a protein, a DNA molecule, or a peptide scaffolding matrix.

9. The method described in claim 6 wherein the second molecule is a macromolecule.

10. The method according to claim 6, wherein the weak base moiety is selected from the group consisting of a 1,2-amino thiol of cysteine, a 1,2-amino ethanol of serine, a 1,2-amino ethanol of threonine, an aminooxyacetyl function, a mono-hydrazine succinyl function, or 4-hyrazino benzoyl.

11. A method of chemically ligating a carboxylic group of a first peptide segment and an amino group of a second peptide segment, said method comprising:

a) introducing an aldehyde ester function to said carboxylic group of said first peptide segment;

b) reacting said function with said amino group, the amino group being ß-functionalized with a weak base moiety.

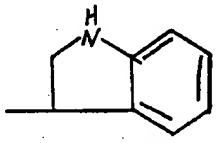
12. The method according to claim 11, wherein said first peptide segment is physically separate and unattached from said second peptide segment.

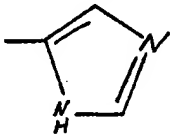
13. The method described in claim 11, wherein said first and said second peptide segments comprise opposite ends of the same polypeptide chain.

14. The method according to claim 11 wherein the carboxylic group is at the C-terminus of the first peptide and the amino group is at the N-terminus of the second peptide.

15. The method according to claim 11, wherein the aldehyde ester function is introduced by esterifying an α-hydroxyaldehyde to said carboxylic group of said first peptide segment.

16. The method according to claim 11, wherein the weak base moiety is selected from the group consisting of amino, sulfhydryl, hydroxy, carboxamide

indole  ,

imidiazole  ,

an aminooxyacetyl function, a mono-hydrazine succinyl function, or 4-hyrazino benzoyl.

17. The method according to claim 11, wherein said aldehyde ester is introduced to said carboxylic group of said first peptide segment by an enzyme-catalyzed coupling reaction.

18. The method according to claim 11, wherein said aldehyde ester function is introduced to said carboxylic group of said first peptide by solid phase synthesis reaction.

19. The method according to claim 11, wherein said aldehyde ester has the formula:

$$-O-(CR_1R_2)n-CHO$$

wherein R_1 R_2 are hydrogen, short chain alkyl or substituted aromatic moieties, and n=1 to 9.

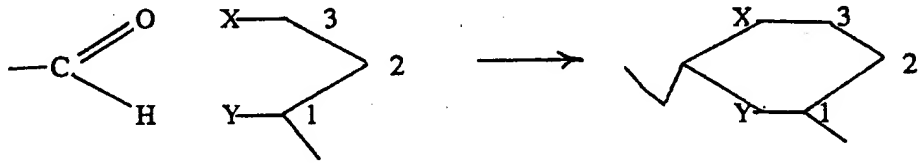
20. The method according to claim 11, wherein said aldehyde ester is an α-formyl methyl ester, an α-formyl ethyl ester or a ß,ß,ß-dimethyl formyl ethyl ester.

21. The method according to claim 11, wherein said method is practiced in a single reaction vessel.

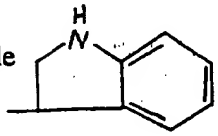
22. The method according to claim 11, further comprising the step of upwardly adjusting the pH of the reaction to enhance an O to N-acyl transfer between the carboxylic and amino groups.

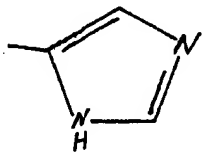
23. The method according to claim 11, wherein the reaction of the weak base moiety and the aldehyde moiety occurs in an aqueous solution.

24. The method according to Claim 11 wherein the carboxylic group is chemically ligated with the amino group by the interaction of the aldehyde with a 1,3-disubstituted α-amino acid which is able to form a 6-member ring having the general structure



where X= $-NH_2$, -SH, -OH, $-\overset{O}{\overset{\|}{C}}NH_2$,

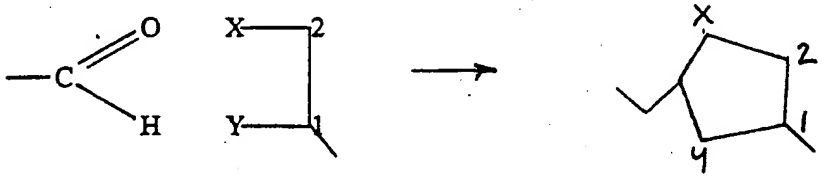
indole 

, imidiazole 

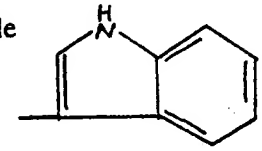
and Y= $-NH_2$, and where the substituent numbers correspond to the numbers set forth in the structures.

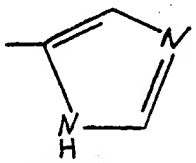
25. The method according to Claim 11 wherein the carboxylic group is

chemically ligated with the amino group by the interaction of the aldehyde with a 1,2-disubstituted α-amino acid which is able to form a 5-member ring having the general structure



where X= -NH_2, -SH, -OH, -$\overset{O}{\overset{\|}{C}}NH_2$,

indole  ,

imidiazole 

and Y= -NH_2, and where the substituent numbers correspond to the numbers set forth in the structures.

26. A method for synthesizing a peptide dendrimer by chemically ligating an unprotected peptide to a scaffolding matrix, said method comprising:

(a) converting an alcohol moiety at an N-terminal amino acid on a branch of said scaffolding matrix to an aldehyde;

(b) reacting a weak base attached to the N-terminal amino acid of said peptide with said aldehyde to form a bonding structure; and,

(c) allowing said reaction to proceed until all said branches of said scaffolding matrix are occupied by a molecule of said peptide.

27. The method described in claim 26 where said scaffolding matrix is octavalent lysinyl multiple antigen protein.

28. The method described in claim 26 where said scaffolding matrix is tetravalent lysinyl multiple antigen protein.

29. The method described in claim 26 where said N-terminal amino acid of each said branch of said scaffolding matrix comprises serine.

30. The method described in claim 26 where said aldehyde is an α-oxoacyl aldehyde.

31. The method described in claim 26 where said weak base is selected from the group consisting of a 1,2-amino thiol of cysteine, a 1,2-amino ethanol of serine, a 1,2-amino ethanol of threonine, an aminooxyacetyl function, a mono-hydrazine succinyl function, or 4-hyrazino benzoyl.

32. The method described in claim 26 where said bonding structure is a thiazolidine ring, an oxime, a hydrazone, or a phenylhydrazone.

33. The method described in claim 26 where said peptide comprises a 24-amino acid fragment from V3-Loop of gp120 of HIV-1, MN strain.

34. The method described in claim 26 where said peptide comprises a 16-amino acid fragment from V3-Loop of gp120 of HIV-1, MN strain.

35. The method described in claim 26 where said peptide comprises a 20-amino acid fragment derived from a surface protein of feline leukemia virus, VA-20.

36. The method described in claim 26 where said peptide comprises a 10-amino acid fragment from ZP3 glycoprotein.

37. A peptide dendrimer produced by the method described in claim 26.

38. The peptide dendrimer described in claim 37, where said scaffolding matrix is octavalent lysinyl multiple antigen protein.

39. The peptide dendrimer described in claim 37, where said scaffolding matrix is tetravalent lysinyl multiple antigen protein.

40. The peptide dendrimer described in claim 37, where said peptide comprises a 24-amino acid fragment from V3-Loop of gp120 of HIV-1, MN strain.

41. The peptide dendrimer described in claim 37, where said peptide comprises a 16-amino acid fragment from V3-Loop of gp120 of HIV-1, MN strain.

42. The peptide dendrimer described in claim 37, where said peptide comprises a 20-amino acid fragment derived from a surface protein of feline leukemia virus, VA-20.

43. The peptide dendrimer described in claim 37, where said peptide

comprises a 10-amino acid fragment from ZP3 glycoprotein.

44. A method for site-specific modification of a macromolecule, the method comprising the step of reacting a weak base moiety of a second molecule with an aldehyde moiety of the macromolecule.

45. The method according to claim 44 wherein said macromolecule is a protein and said second molecule is a lipid module, said lipid module comprised of a lipid and a peptide, further comprising the steps of:

(a) converting an alcohol moiety of an amino acid of said protein to an aldehyde; and

(b) reacting a weak base moiety of an amino acid of said peptide of said lipid module with said aldehyde.

46. The method described in claim 45, where said amino acid of said protein is at the N-terminus.

47. The method described in claim 45, where said amino acid of said protein is at the C-terminus.

48. The method described in claim 45, where said lipid module is tripalmitoyl glyceryl cysteine.

49. The method described in claim 45, where said amino acid of said peptide of said lipid module is cysteine.

50. The method described in claim 45, where said protein is recombinant derived gp120 of HIV-1, IIIb strain.

51. The method described in claim 46, where said N-terminal amino acid of said protein is serine.

52. The method described in claim 46, where said N-terminal amino acid of said protein is threonine.

53. The method described in claim 49, where said weak base moiety comprises 1,2-amino thiol of cysteine.

54. The method described in claim 52, where said aldehyde is a glyoxylyl aldehyde.

55. A method for site-specific modification of a macromolecule, the method comprising the step of reacting an aldehyde moiety of a second molecule with a weak

base moiety of the macromolecule.

56. The method according to claim 55 wherein the macromolecule is a protein.

57. The method described in claim 56 where the second molecule is a lipid module.

58. The method described in claim 56 where the second molecule is MeO-PEG-CHO.

59. The method described in claim 56 where said protein is recombinant gp 120.

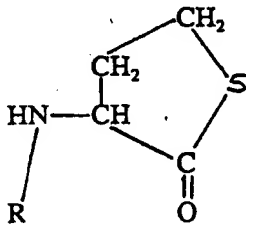
60. The method described in claim 56 where said protein is interleukin-2.

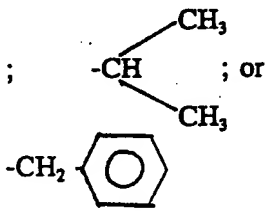
61. The method described in claim 56 further comprising the steps of:

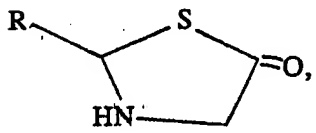
(a) introducing a 1,2- or 1,3-amino thiol ring moiety by reacting a secondary amine thiolactone with the protein; and,

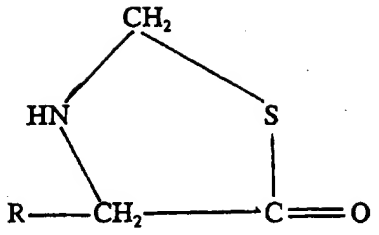
(b) reacting said thiol moiety with an aldehyde of the second molecule.

62. The method described in claim 61 where said secondary amine thiolactone is selected from the group consisting of:



, where R= -CH_3 ;  ; or



or 

63. A macromolecule modified by the method described in Claim 44.

64. A macromolecule modified by the method described in Claim 55.

65. A method for synthesizing a vaccine by chemically ligating a peptide to a macromolecule, said method comprising the steps of reacting a weak base moiety of the peptide with an aldehyde of the macromolecule.

66. The method described in claim 65 wherein the macromolecule is a protein or a peptide scaffolding matrix.

67. A method for synthesizing a vaccine by chemically ligating a peptide to a macromolecule, said method comprising reacting an aldehyde moiety of the peptide with a weak base moiety of the macromolecule.

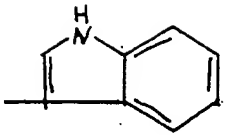
68. The method described in claim 67 wherein the macromolecule is a protein or a peptide scaffolding matrix.

69. A method of chemically ligating a first macromolecule to a second molecule, said method comprising reacting two weak base moieties, one weak base moiety being ß-functionalized with the other weak base moiety, on the first macromolecule with an aldehyde moiety on the second molecule.

70. The method described in claim 69 wherein said reaction of the weak base moieties with the aldehyde moiety is performed under acidic conditions.

71. The method described in claim 69 wherein one of the weak base moieties is an amino group.

72. The method according to claim 71 wherein the other weak base moiety is selected from the group consisting of amino, sulfhydryl, hydroxy, carboxamide - $\overset{O}{\overset{\|}{C}}NH_2$,

indole  , or

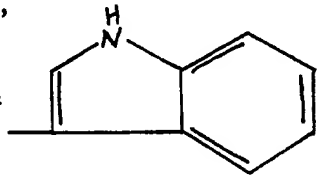
imidiazole 

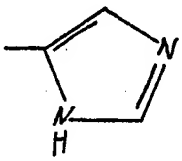
73. A method of chemically ligating a first macromolecule to a second molecule, said method comprising reacting two weak base moieties, one weak base moiety being ß-functionalized with the other weak base moiety, on the second molecule with an aldehyde moiety on the first macromolecule.

74. The method described in claim 73 wherein said reaction of the weak base moieties with the aldehyde moiety is performed under acidic conditions.

75. The method described in claim 73 wherein one of the weak base moieties is an amino group.

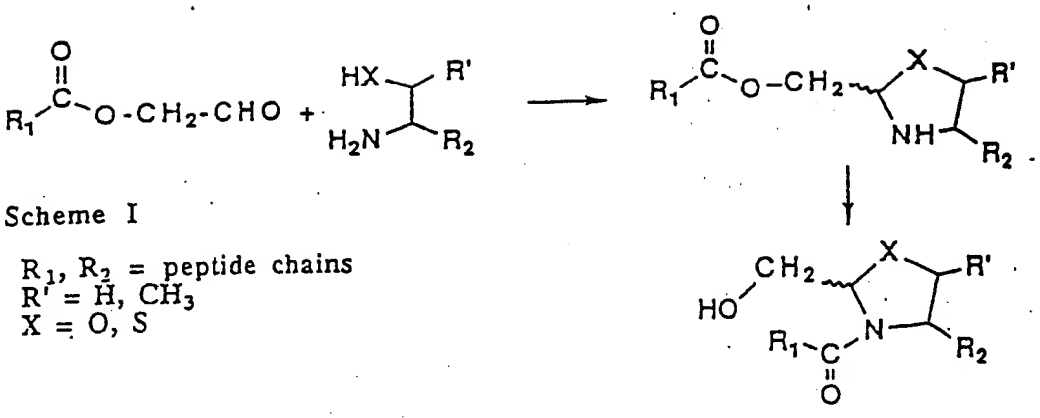
76. The method according to claim 75 wherein the other weak base moiety is selected from the group consisting of amino, sulfhydryl, hydroxy, carboxamide -  ,

indole  , or

imidiazole 

STATEMENT UNDER ARTICLE 19

Pursuant to Article 19 of the Patent Cooperation Treaty and Rule 46, Applicant respectfully submits the attached sheets of amended claims. These sheets contain new claims 1-10, 14, 16-18, 20, 44, 46, 47, 55, 56, and 65-76. The new claims do not go beyond the scope of the application as filed. Claims numbered 1-16 in the application as filed have been amended and rewritten as claims 11-25. Claims numbered 17-27 in the application as filed have been renumbered to claims 26-36. Claims numbered 28-34 in the application as filed have been renumbered to claims 37-43. Claims numbered 35-42 in the application as filed have been amended and renumbered as claims 45 and 48-54. Claims numbered 43-48 in the application as filed have been amended and rewritten as claims 57-62. Claims numbered 49 and 50 in the application as filed have been rewritten as claims 63 and 64.

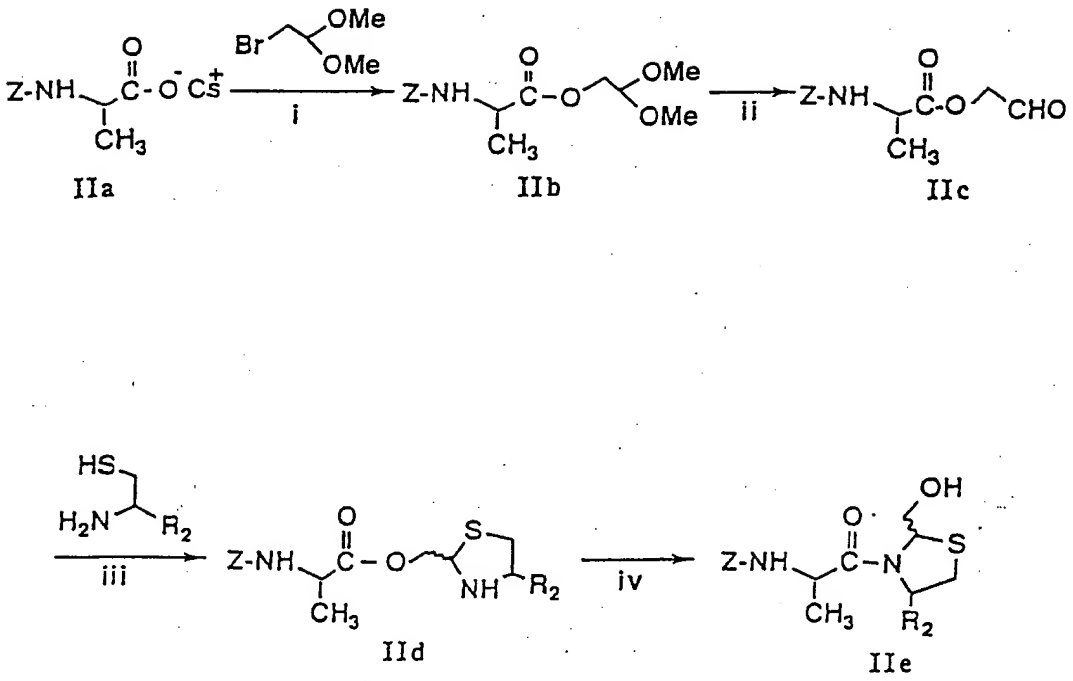


Scheme I

R_1, R_2 = peptide chains
R' = H, CH_3
X = O, S

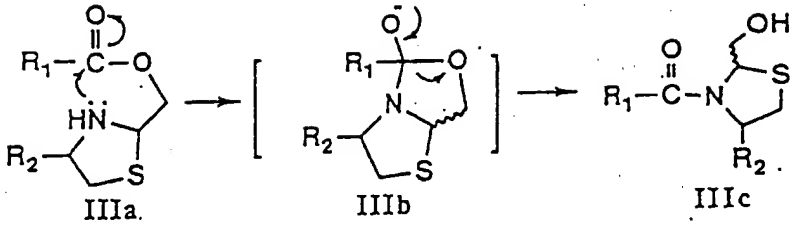
FIGURE 1

FIGURE 2



Scheme II.

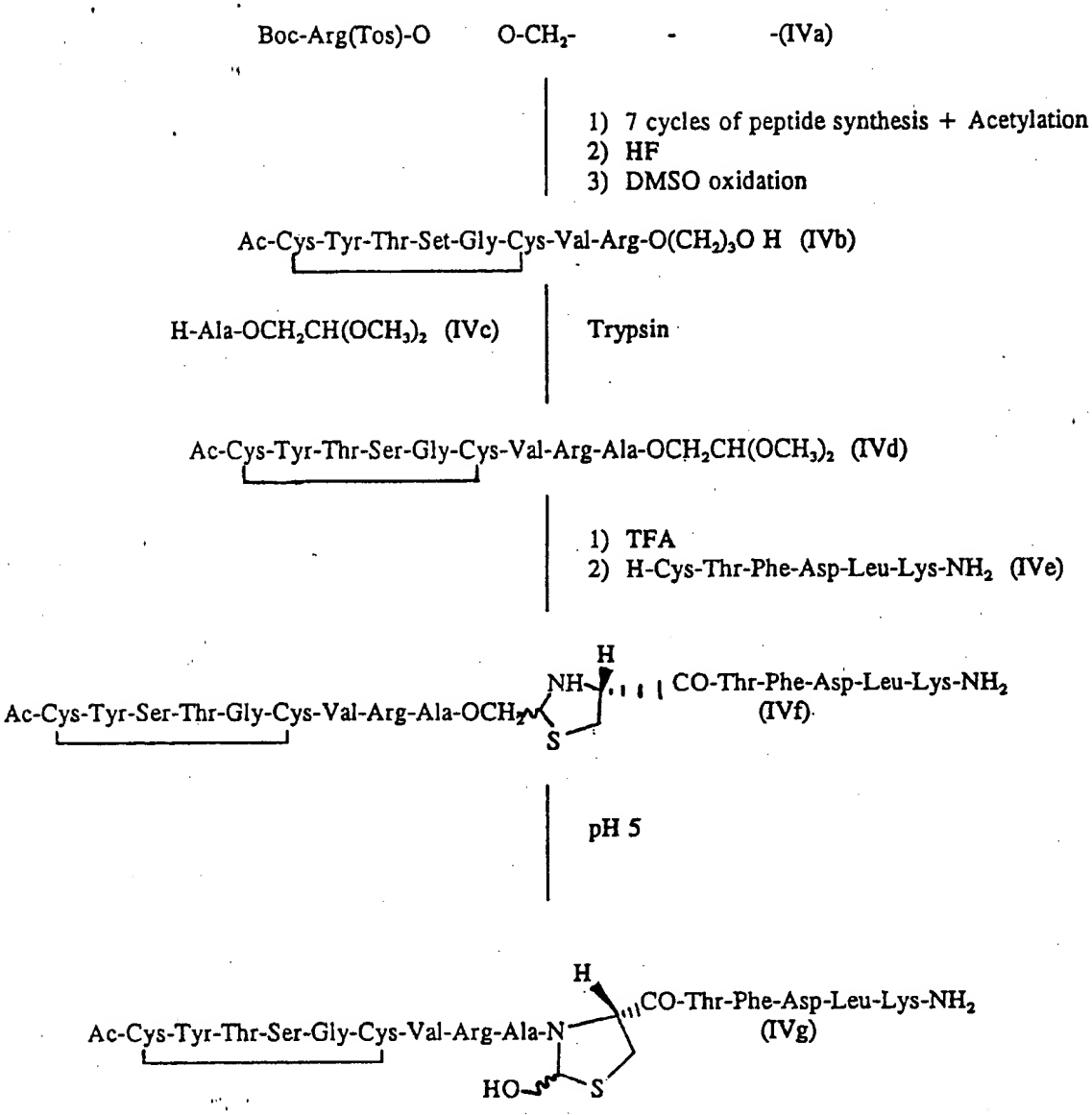
i. DMF, 60-70°C, 24 h; ii. 30% TFA in CH_2Cl_2 (2-5% H_2O); iii. H_2O/CH_3CN, pH2-5; iv. pH6-9.



Scheme III. Mechanism of the acyl transfer reaction.

FIGURE 3

FIGURE 4



Scheme IV. Synthesis of a pentadecapeptide.

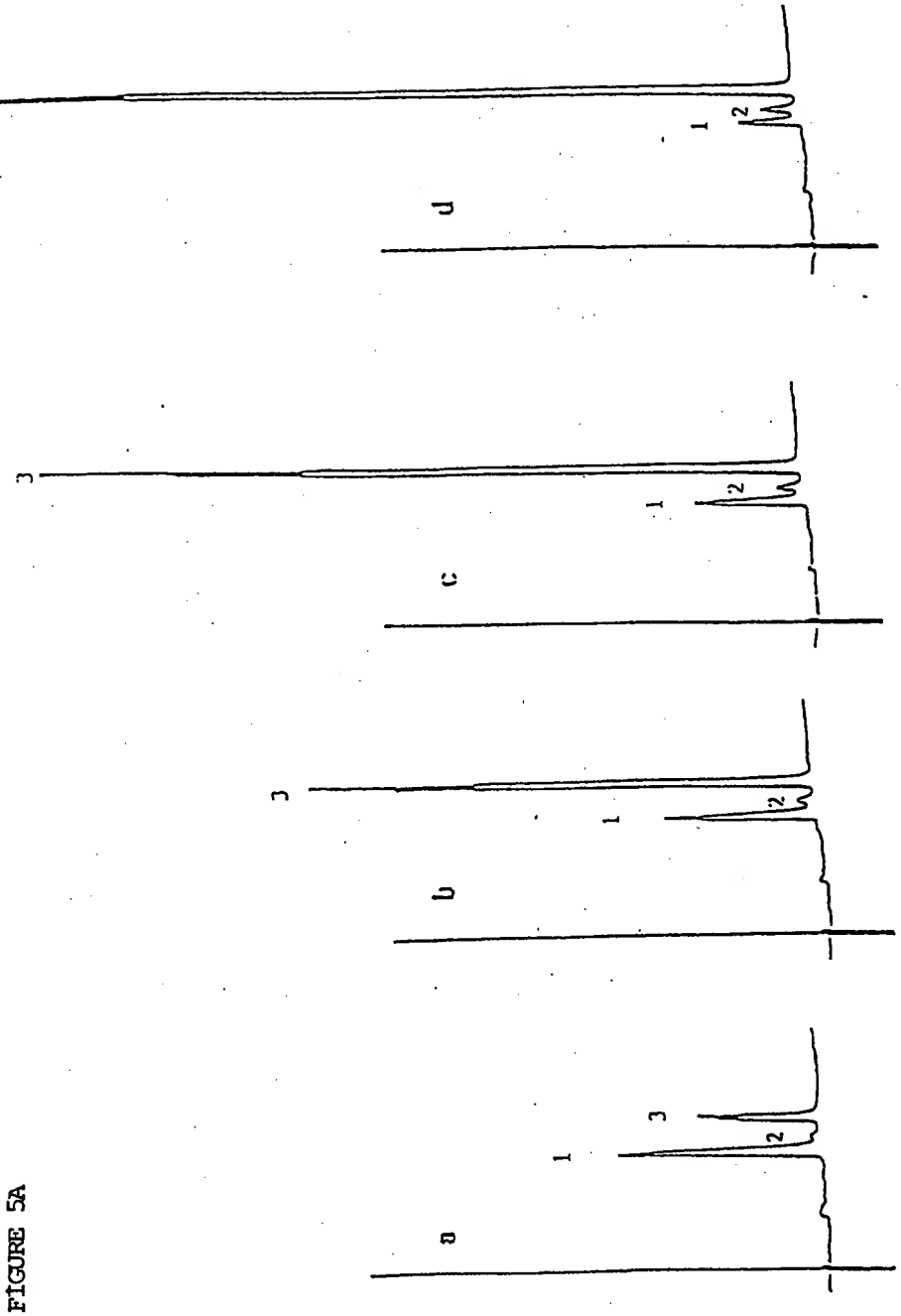


FIGURE 5A

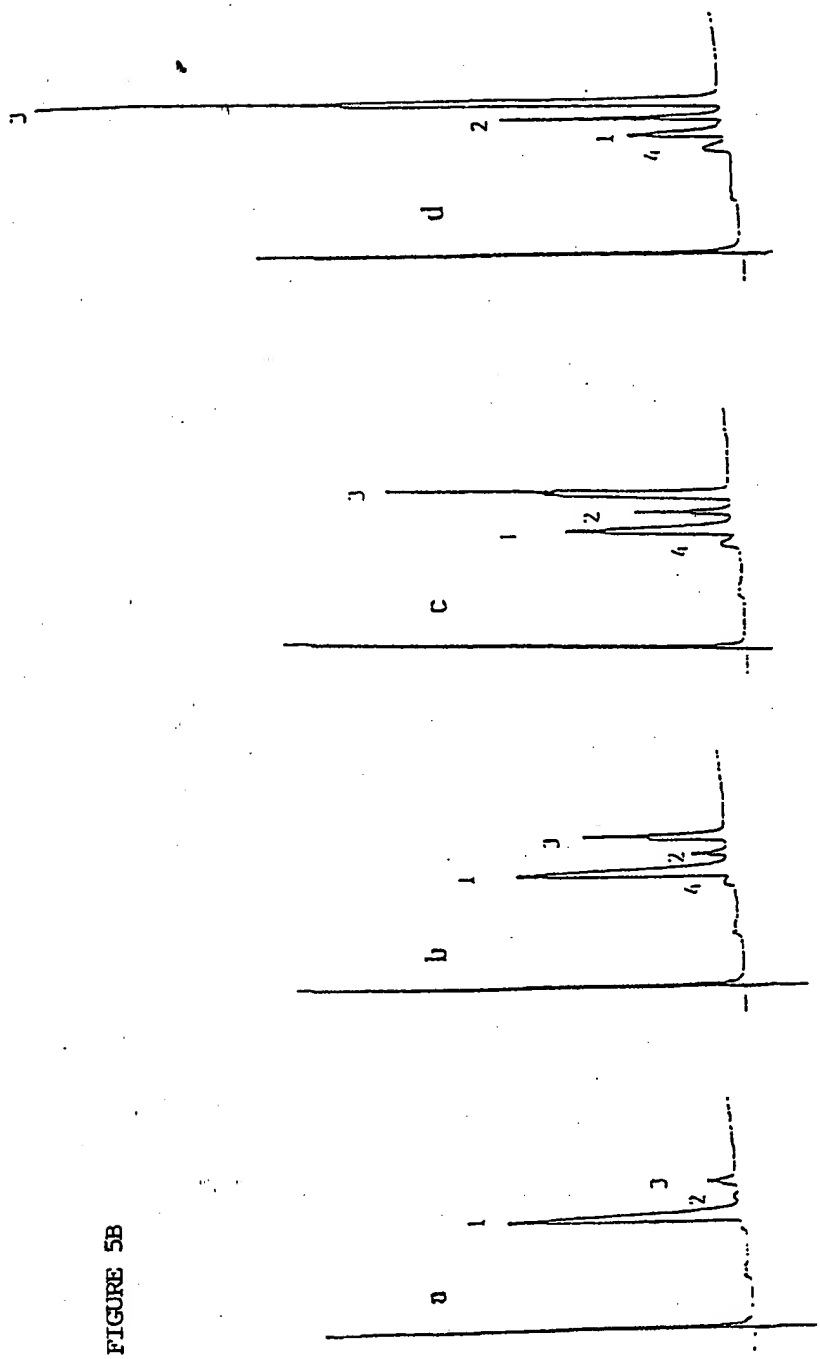


FIGURE 5B

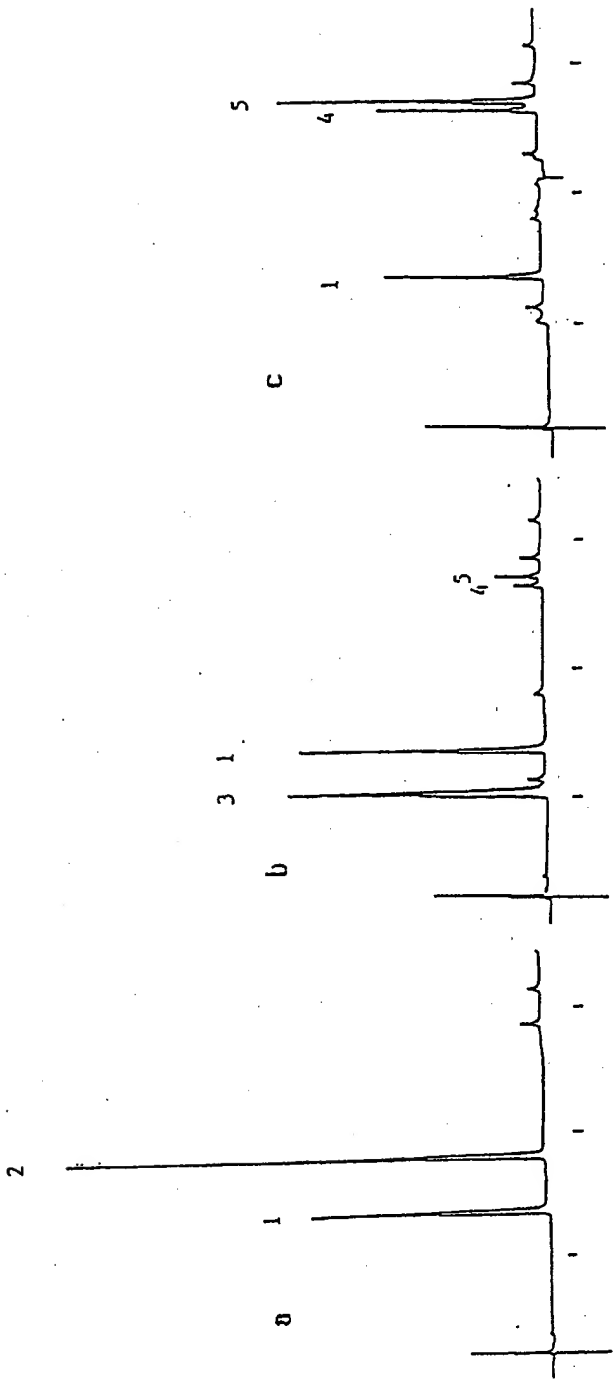


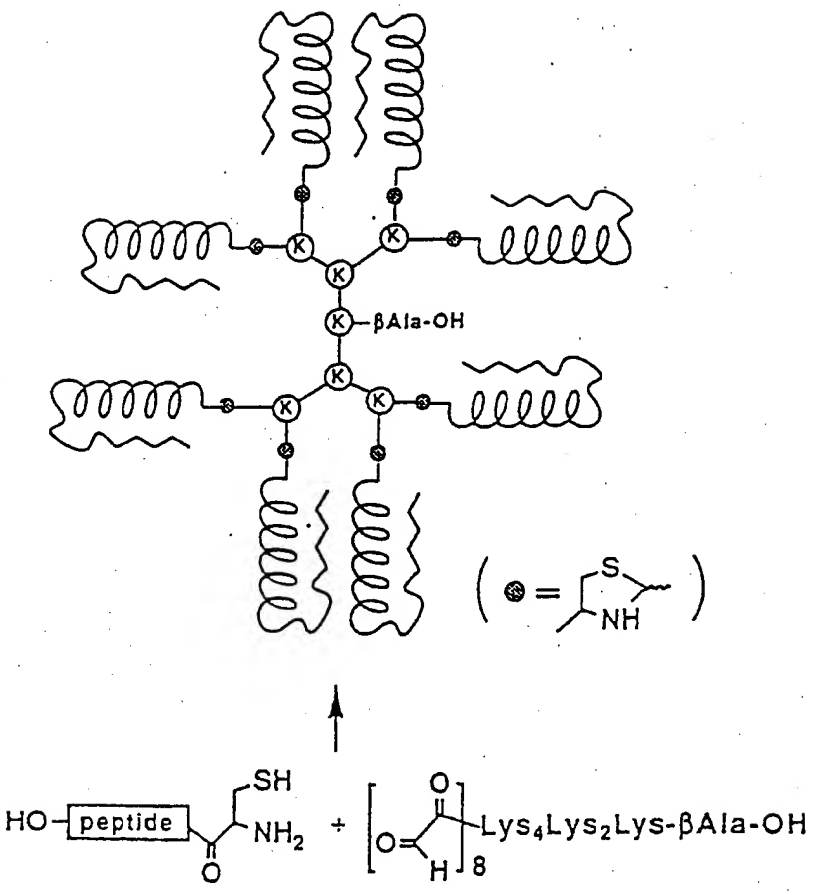
FIGURE 6

FIGURE 7

Table 1. Rate ($t_{1/2}$, hr) of the O- to N- Acyl Transfer Reaction

pH	5	6	7	7.4	8	9
IId1 to IIe1		37.5	22.2	20.2	9.9	3.4
IId2 to IIe2		55	8.4	9.5	11	9.3

FIGURE 8



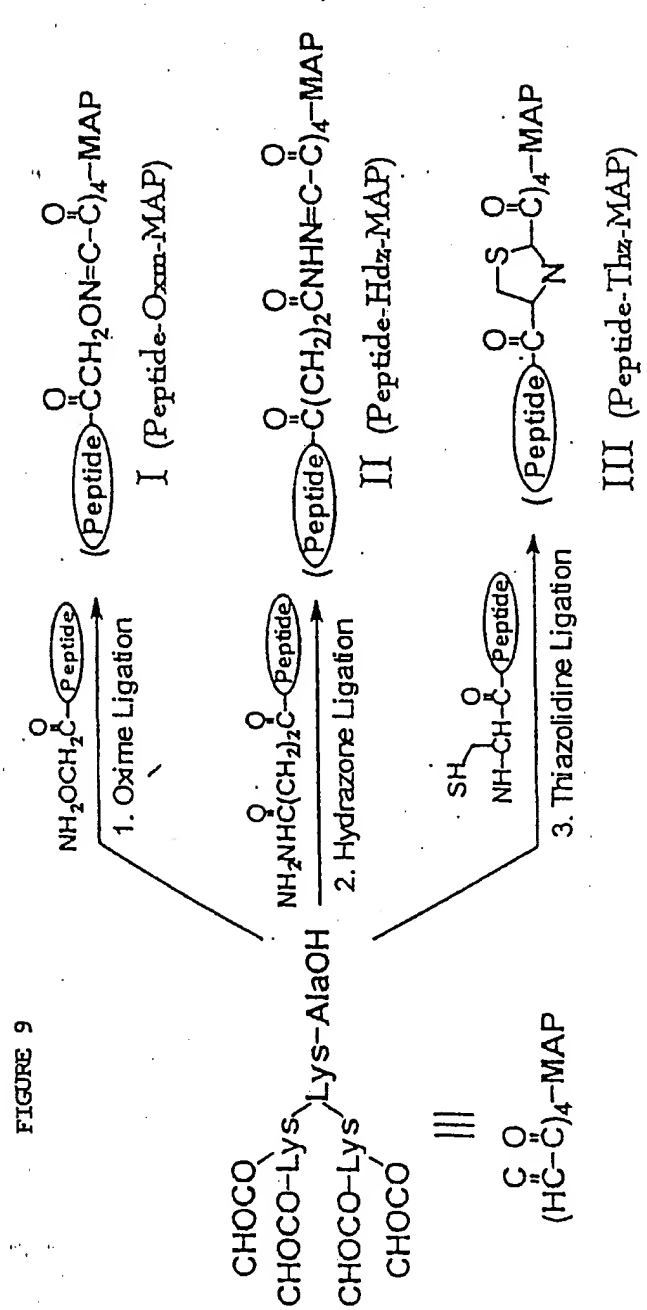

1. Oxime Ligation
2. Hydrazone Ligation
3. Thiazolidine Ligation
I (Peptide-Oxm-MAP)
II (Peptide-Hdz-MAP)
III (Peptide-Thz-MAP)
Lys-AlaOH
MAP

FIGURE 9

FIGURE 10

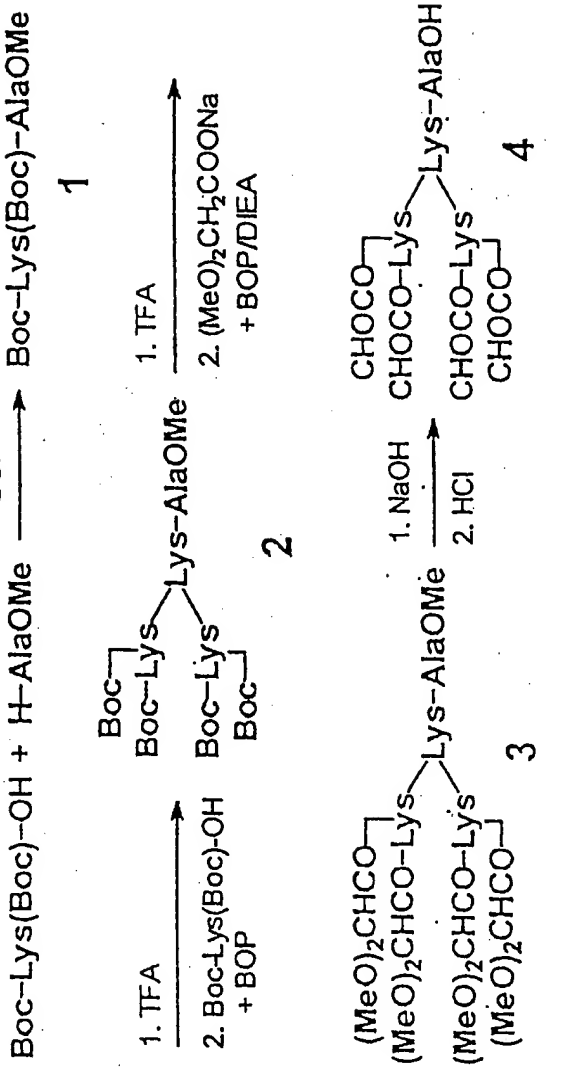

Boc–Lys(Boc)–OH + H–AlaOMe
BOP
Boc–Lys(Boc)–AlaOMe
1
1. TFA
2. Boc-Lys(Boc)-OH + BOP
Boc
Boc–Lys
Boc–Lys
Boc
Lys–AlaOMe
2
1. TFA
2. (MeO)2CH2COONa + BOP/DIEA
(MeO)2CHCO
(MeO)2CHCO–Lys
(MeO)2CHCO–Lys
(MeO)2CHCO
Lys–AlaOMe
3
1. NaOH
2. HCl
CHOCO
CHOCO–Lys
CHOCO–Lys
CHOCO
Lys–AlaOH
4

FIGURE 11

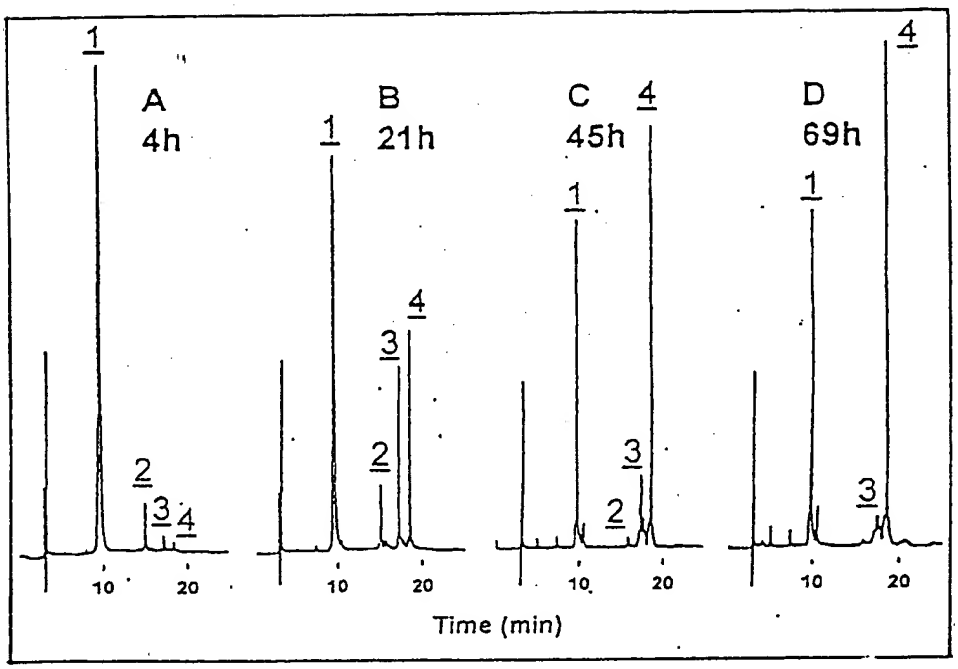


FIGURE 12

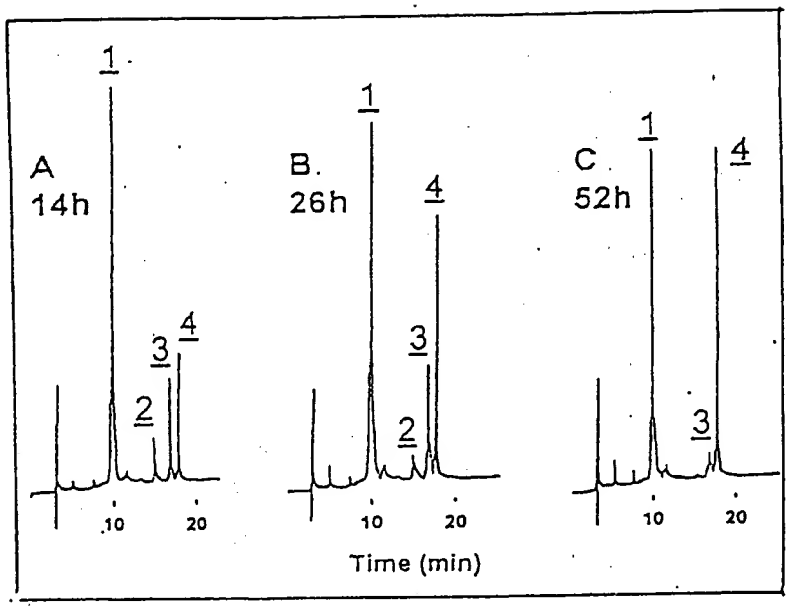


FIGURE 13

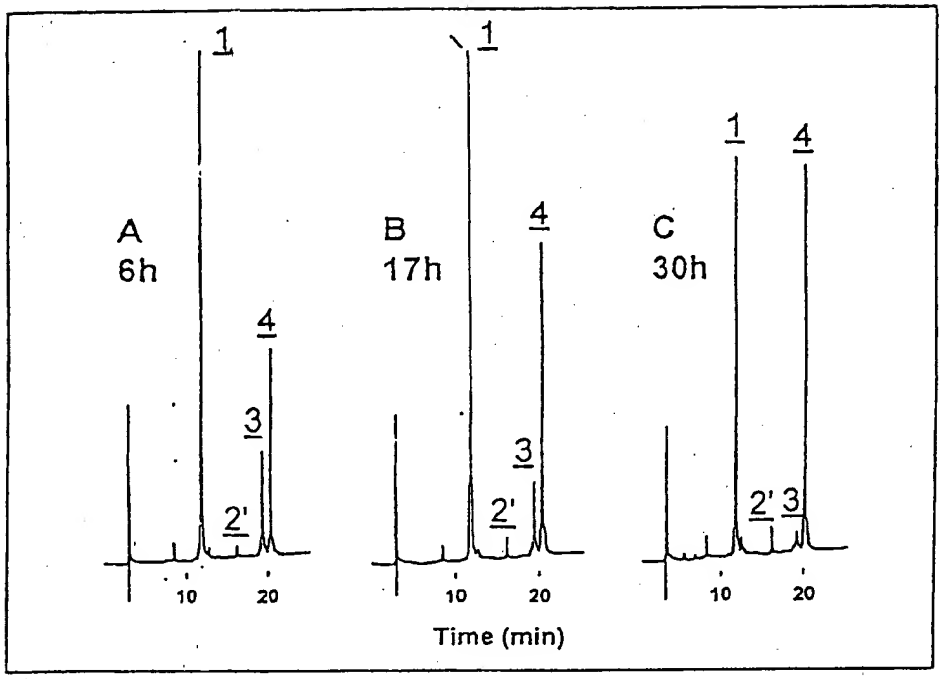


FIGURE 14A

Table 1. Rate of VA-20-MAP formation through oxime ligation

Reaction Condition	pH				50% Organic cosolvents (pH 5.7)			37 °C			
								H_2O		50% DMSO	
	4.7[3]	4.2	5.2	5.7	CH_3CN	DMF	DMSO	pH 4.7	pH 4.7	pH 5.7	
Time (h)[1]	52	64	38	32	35	18	8	23	16	4.5	
Rel. Rate[2]	1.0	0.8	1.4	1.6	1.5	2.9	6.5	2.3	3.3	12	

1: time for reaching 90% of completion of ligation reaction based on HPLC analysis
2: Rel. Rate: relative reaction rate based on standard condition.
3: defined as standard condition.

FIGURE 14B

Table 2. Rate of VA-20-MAP ligation through hydrazone ligation

Reaction condition	pH			50% Organic solvents (pH 5.7)			37 °C			
							H_2O	50%DMSO		
	5.2[3]	4.7	5.7	CH_3CN	DMF	DMSO	pH 5.2	pH 5.2	pH 5.7	
Time (h)[1]	40	44	34	76	16	2	26	8	1.5	
Rel. Rate[2]	1.0	0.9	1.2	0.5	2.5	20	1.5	5	27	

1: time for reaching 90% of completion of ligation reaction based on HPLC analysis
2: Rel. Rate: relative reaction rate based on standard condition.
3: defined as standard condition.

FIGURE 14C

Table 3. Rate of VA-20-MAP ligation through thiazolidine ligation

Reaction condition	pH			50% Organic solvents (pH 4.5)			37 °C	
							H_2O	50% DMF
	4.5[3]	4.0	5.0	CH_3CN	TFE	DMF	pH 4.5	pH 4.5
Time (h)[1]	24	30	16	18	30	5	8	2
Rel. Rate[2]	1.0	0.8	1.5	1.3	0.8	4.8	3	12

[1]: time for reaching 90% of completion of ligation reaction based on HPLC analysis
[2]: Rel. Rate: relative reaction rate based on standard condition.
[3]: defined as standard condition.

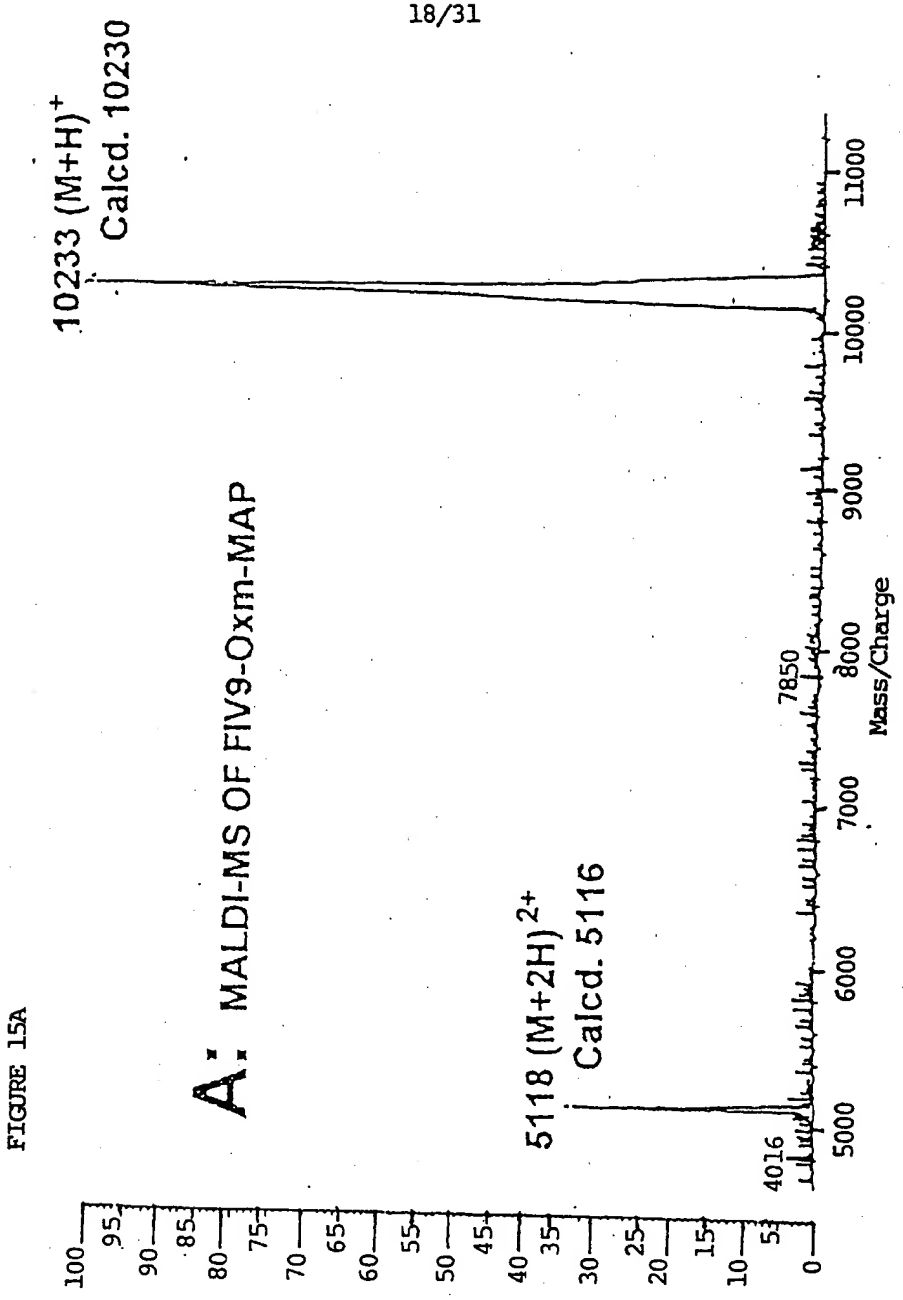

A: MALDI-MS OF FIV9-Oxm-MAP
10233 (M+H)+
Calcd. 10230
5118 (M+2H)2+
Calcd. 5116
4016
7850
Mass/Charge
100
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
5000
6000
7000
8000
9000
10000
11000

FIGURE 15A

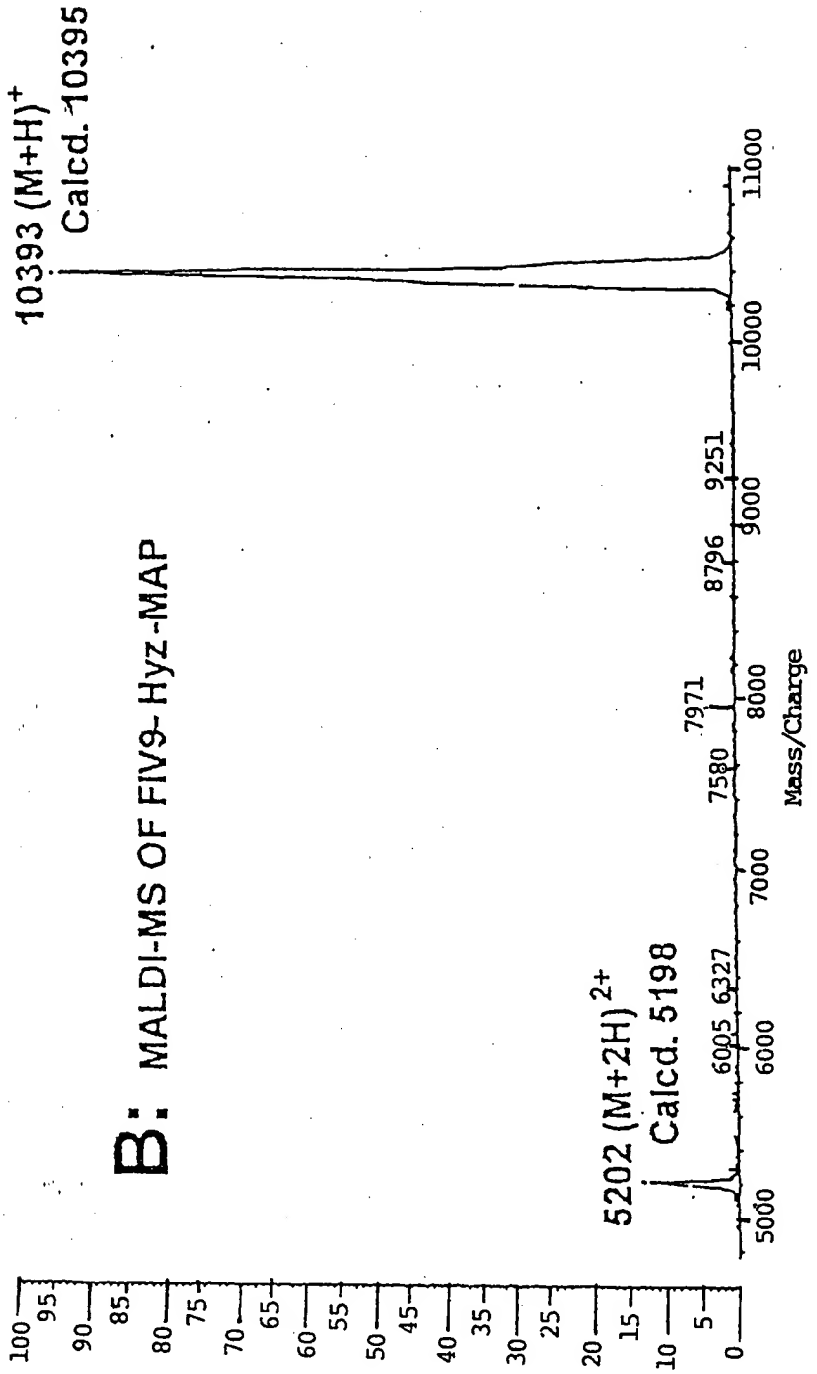

B: MALDI-MS OF FIV9- Hyz-MAP
10393 (M+H)+
Calcd. 10395
5202 (M+2H)2+
Calcd. 5198
Mass/Charge
5000
6000
7000
8000
9000
10000
11000
6005
6327
7580
7971
8796
9251
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100

FIGURE 15B

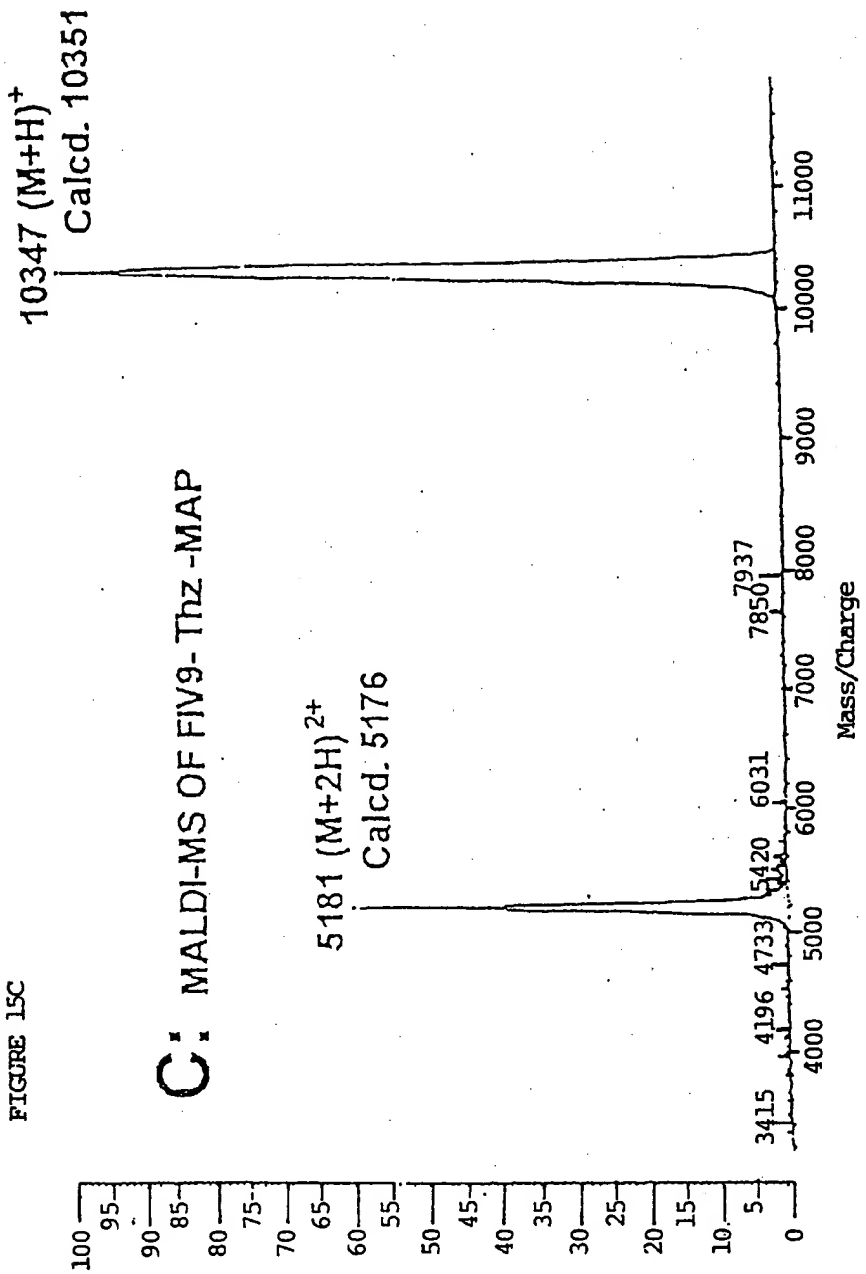

C: MALDI-MS OF FIV9- Thz -MAP
10347 (M+H)+
Calcd. 10351
5181 (M+2H)2+
Calcd. 5176
3415
4196
4733
5420
6031
7850
7937
Mass/Charge
100
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
4000
5000
6000
7000
8000
9000
10000
11000

FIGURE 15C

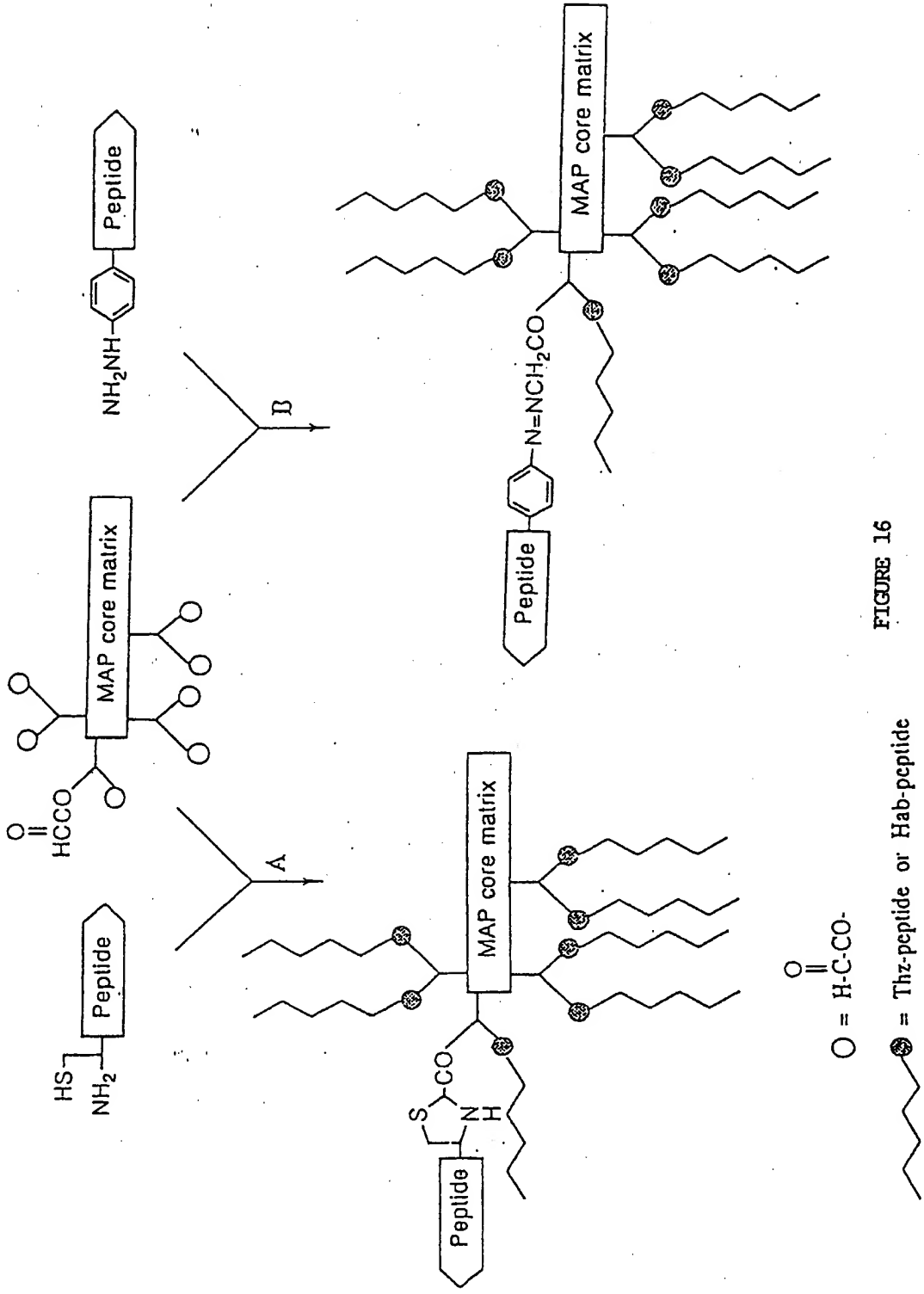

MAP core matrix
Peptide
HS
NH2
HCCO
A
B
NH2NH
N=NCH2CO
S
CO
NH
O = H-C-CO-
= Thz-peptide or Hab-peptide

FIGURE 16

FIGURE 17

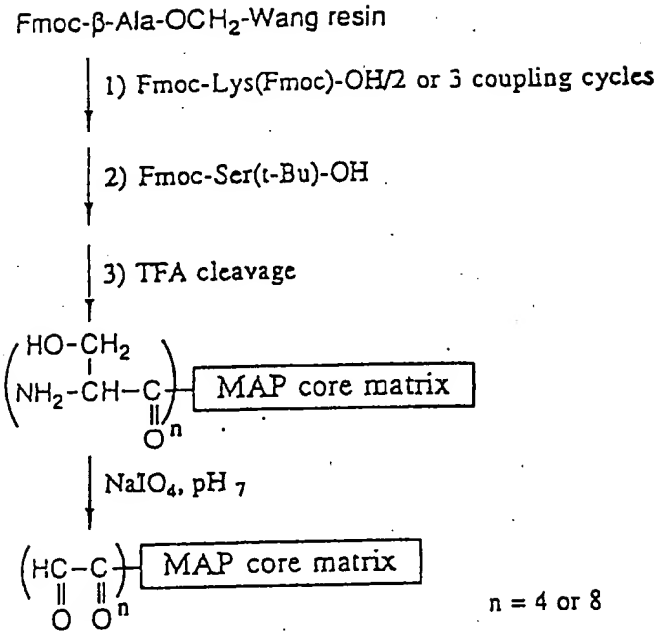

Fmoc-β-Ala-OCH2-Wang resin
1) Fmoc-Lys(Fmoc)-OH/2 or 3 coupling cycles
2) Fmoc-Ser(t-Bu)-OH
3) TFA cleavage
HO-CH2
NH2-CH-C
MAP core matrix
n
NaIO4, pH 7
HC-C
MAP core matrix
O O n
n = 4 or 8

FIGURE 18

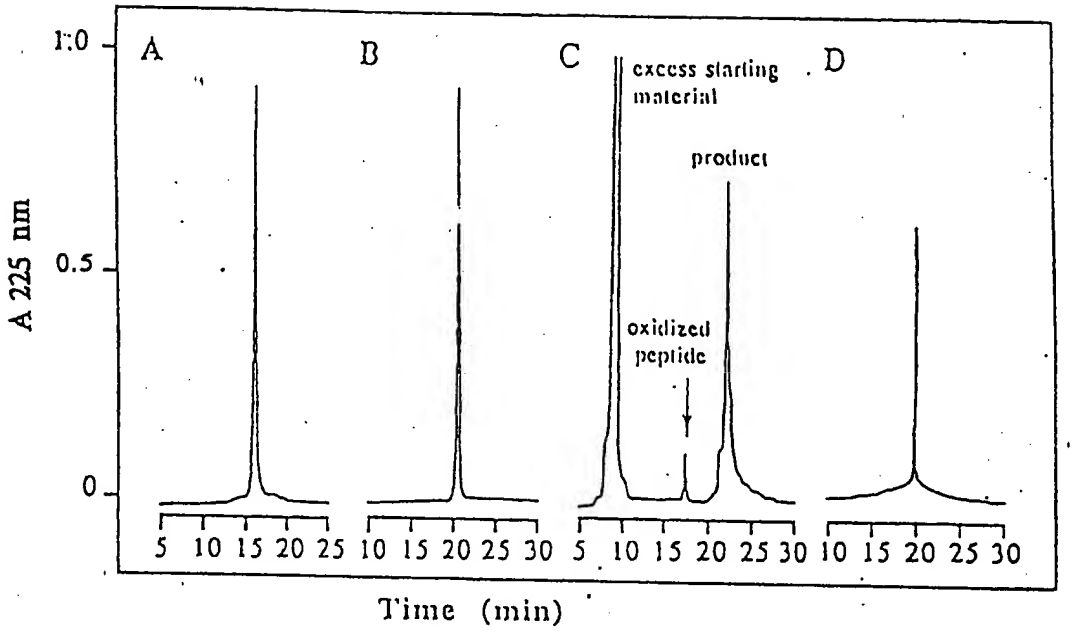

1.0
0.5
0
A 225 nm
A
B
C
D
excess starting material
product
oxidized peptide
5 10 15 20 25
10 15 20 25 30
5 10 15 20 25 30
10 15 20 25 30
Time (min)

FIGURE 19

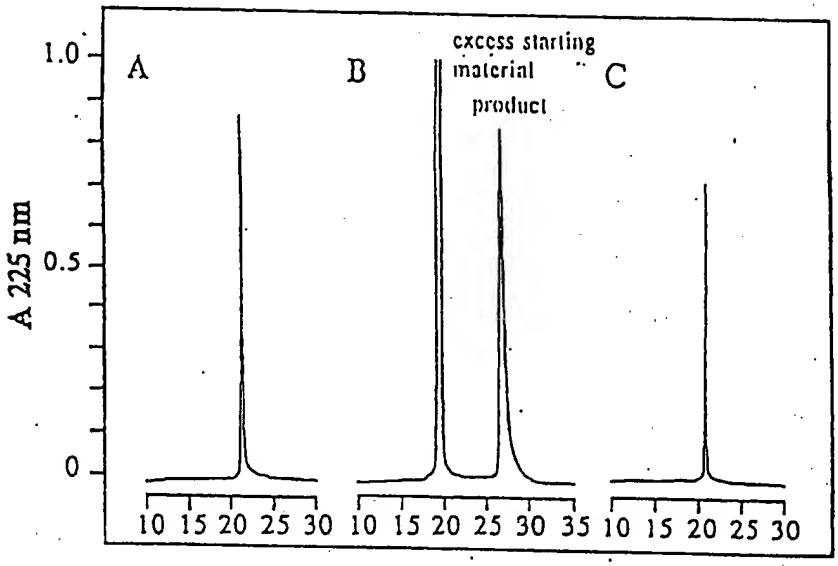

A
B
C
excess starting material
product
A 225 nm
1.0
0.5
0
10 15 20 25 30
10 15 20 25 30 35
10 15 20 25 30

FIGURE 20

Table 1

Effect of antioxidant, organic solvent and temperature on the thiazolidine ring on the ligation of anunprotected peptide CA-16 to $(CHO)_4$-MAP.

Reaction conditions[1]	Conc. peptide (mM)	Time (h) for 50% completion 22 °C	37 °C	50 °C
H_2O	4.8	15	-	
H_2O + EDTA	4.8	10.7	5.8	2
H_2O + EDTA	8.0	6.3	5	2.5
H_2O/DMF (4:6, v/v)	4.8	4.5	2	1.6

[1] All the reactions were performed at pH 5 using 4 mol equiv. of peptide (CA-16) for each α-oxoacyl group on $(CHO)_4$-MAP and the pseudo first order rates were used.

FIGURE 21

Table 2

Analysis of peptide fragments and ligation products by matrix-assisted laser desorption mass spectrometry

compound	MW from $(M + 1)^+$	calcd values	Δ
$(Ser)_4$-MAP	822	821	+1.0
$(Ser)_8$-MAP	1,683	1,682.5	+0.5
Hob-SR10	1,291	1,290.5	+0.5
CA16	1,884	1,884	0.0
$(SR10)_4$-Hab-MAP	5,789.1	5,788	+1.1
$(SR10)_8$-Hab-MAP	11,615	11,614.5	+0.5
$(NA15)_4$-Thz-MAP	8,162	8,160.7	+1.3
$(NA15)_8$-Thz-MAP	16,363	16,362.5	+0.5

FIGURE 22

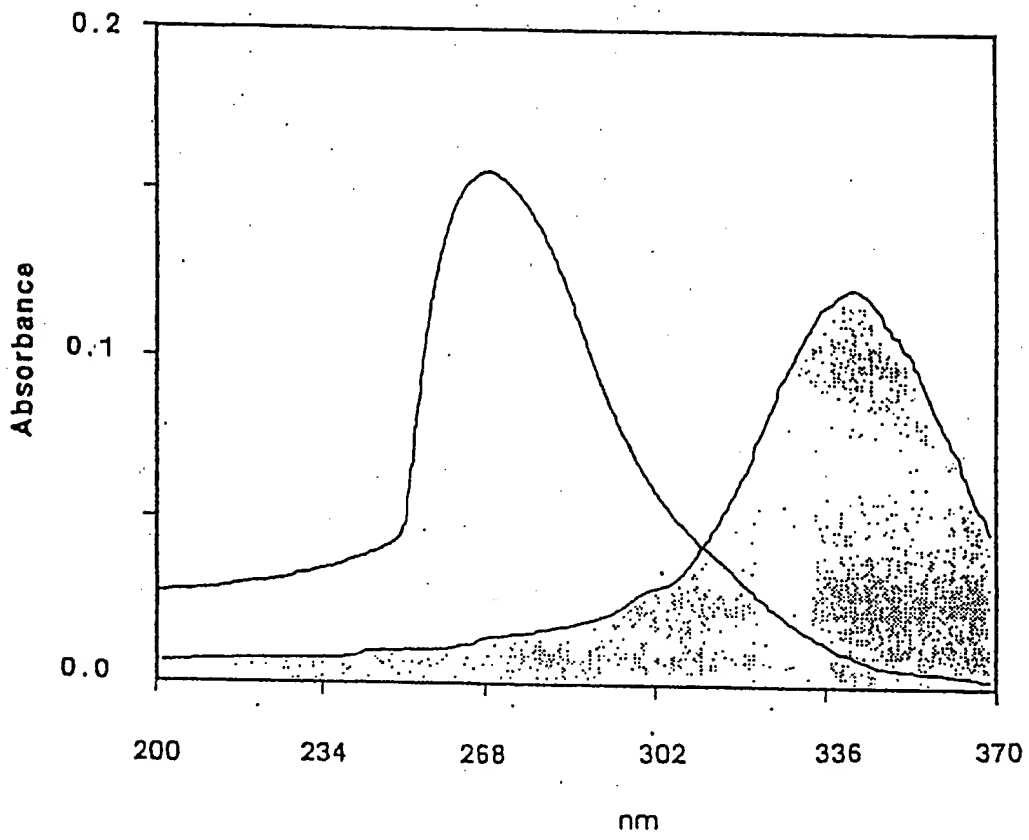
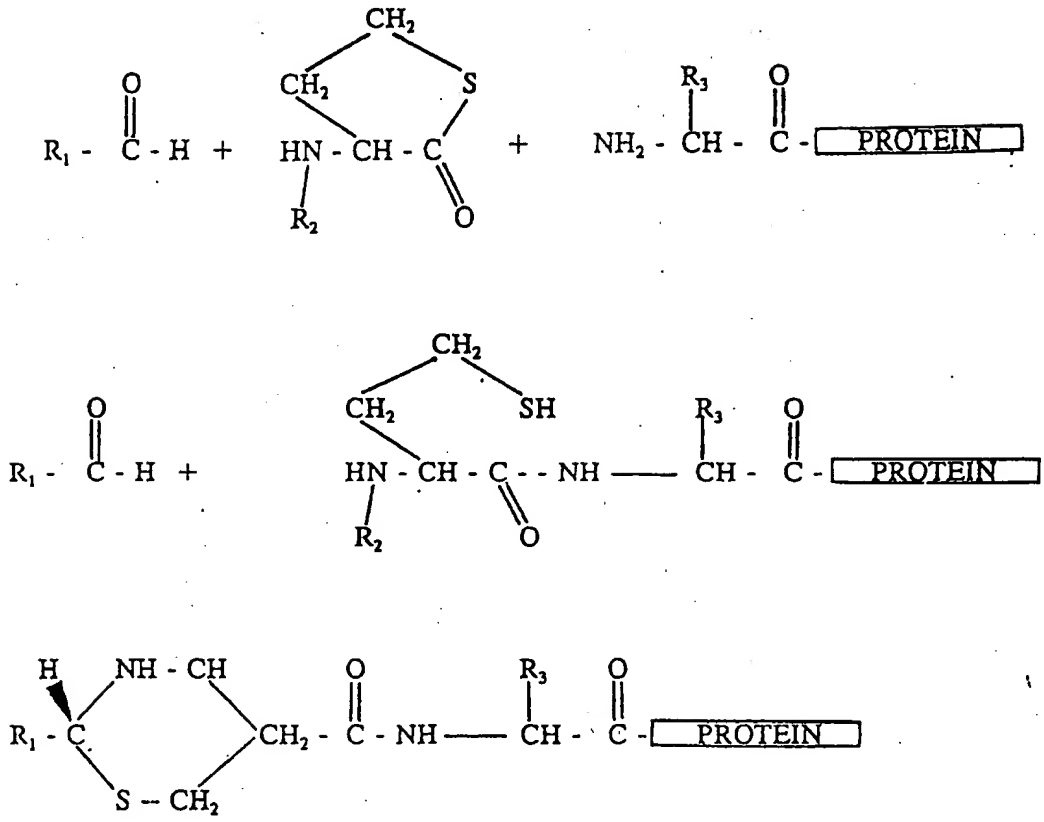


Figure 23A

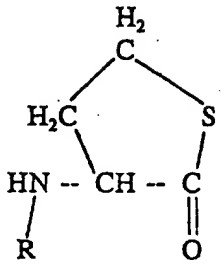
REACTION SCHEME AND THIOLACTONES

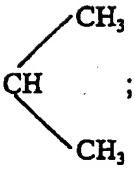


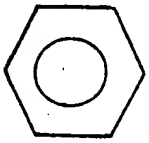
SUBSTITUTE SHEET (RULE 26)

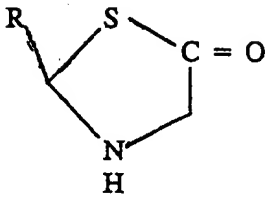
Figure 23B

SUBSTITUTED THIOLACTONES

1. 

FOR NOS. 1, 2, 3: R = - CH_3 ; -  ;

- CH_2 - 

2. 

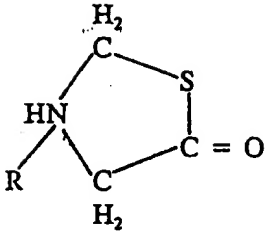
3. 

FIGURE 24

Abbreviations:

Ser	Serine	Z-Ala	Carbobenzoxyalanine
His	Histidine	FM	Formylmethyl Ester
Thr	Threonine	TFA	Trifluoroacetic Acid
Tyr	Tyrosine	HPLC	High Pressure Liquid Chromatography
Cys	Cysteine	NaOH	Sodium Hydroxide
Lys	Lysine	NMR	Nuclear Magnetic Resonance
Arg	Arginine	Z-Ala-Pro	Carbobenzoxyalalyl Proline
Pro	Proline	X	Any Amino Acid
Ile	Isoleucine	ACM	Acetamidomethyl
Phe	Phenylalanine	Ala ODMOE	2-dimethoxyethyl Ester of Alanine
Trp	Tryptophan	DMSO	Dimethylsulfoxide
Leu	Leucine	DMF	Dimethylformamide
Val	Valine	Ala	Alanine
Gly	Glycine		

FIGURE 25

Boc	t-butyloxycarbonyl
Bzl	benzyl
DCC	dicyclohexylcarbodiimide
DCM	dichloromethane
DIEA	diisopropylethyl amine
Dnp	dinitrophenyl
EDTA	ethylenediaminetetraacetic acid
Fmoc	fluorenylmethyloxycarbonyl
Hab	aldehyde-4-hydrazinobenzoyl adduct
HBTU	2-(1H-benxotriazol-1-yl)-1,1,3,3-tetramethyluronium
Hob	4-hydrazinobenzoyl
HOBt	1-hydroxybenzotriazole
LDMS	matrix assisted laser-desorption mass spectrometry
Thz	4-hydroxymethyl riazolidinyl carboxylic acid
Tos	toluenesulphonyl
Z	benzyloxycarbonyl

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/07222

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :G01N 33/531; C07K 1/02, 1/10, 3/06, 3/08, 5/02, 7/02
US CL : 436/543; 530/333, 338, 339, 345, 402
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 436/543; 530/333, 338, 339, 345, 402

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CAS, REGISTRY, APS
glcyoaldehyde, peptide, segment condensation or synthesis

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ANGEWANTE CHEMIE INT. ED. ENGL., Volume 30, Number 2, issued February 1991, E. Bayer, "Towards the Chemical Synthesis of Proteins", pages 113-129.	1-27
A	ANN. REV. BIOCHEM., Volume 57, issued 1988, S.B.H. Kent, "Chemical Synthesis of Peptides and Proteins", pages 957-989.	1-27
A	J. AM. CHEM. SOC., Volume 111, issued 1989, H. Kitaguchi et al., "Enzymatic Peptide Synthesis Via Segment Condensation in the Presence of Water Mimics", pages 9272-9273.	1-27

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 12 SEPTEMBER 1994	Date of mailing of the international search report SEP 22 1994
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer JON P. WEBER, PH.D. Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/07222

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SCIENCE, Volume 256, issued 10 April 1992, M. Schnolzer et al., "Constructing Proteins by Dovetailing Unprotected Synthetic Peptides: Backbone-Engineered HIV Protease", pages 221-225.	1-27
A	J. AM. CHEM. SOC., Volume 93, Number 11, issued 02 June 1971, R. Hirschmann et al., "The Controlled Synthesis of Peptides in Aqueous Medium. VIII. The Preparation and Use of Novel alpha-Amino Acid N-Carboxyanhydrides", pages 2746-2754.	1-27
A	SCIENCE, Volume 243, issued 13 January 1989, E.T. Kaiser et al., "Peptide and Protein Synthesis by Segment Synthesis-Condensation", pages 187-192.	1-27
A	TETRAHEDRON LETT., Volume 31, Number 15, issued 1990, S. Kwiatkowski et al., "Thiazolidine and Thiazoline Derivatives of 3-Aryl-3-Trifluoromethyldiazirines for the Preparation of Fluorescent or 35-S-Radiolabeled Photoaffinity Probes", pages 2093-2096.	1-27
A	J. ORG. CHEM., Volume 54, Number 15, issued 1989, D.S. Kemp et al., "Boc-L-Dmt-OH as a Fully N,S-Blocked Cysteine Derivative for Peptide Synthesis by Prior Thiol Capture. Facile Conversion of N-Terminal Boc-L-Dmt-Peptides to H-Cys(Scm)-Peptides", pages 3640-3646.	1-27, 43-48
A	LIEBIGS ANN. CHEM. Volume 5, issued 1979, R. Bogner et al., "Die Umsetzung von al-D-Galaktosederivaten mit L-Cystein. Beitrage zur Stereochemie von 2-(Polydroxyalkyl)thiazolidin-4-carbonsauren", pages 1637-1657.	1-27, 43-48
A	US, A, 5,118,810 (SANTINI) 02 June 1992.	1-27
A,P	US, A, 5,304,631 (STEWART ET AL.) 19 April 1994.	1-27
A	INT. J. PEPTIDE PROTEIN RES., Volume 27, issue 1986, J. Blake, "Total Synthesis of S-Carbamoylmethyl Bovine Apocytochrome C by Segment Condensation", pages 191-200.	1-27
A	J. ORG. CHEM., Volume 54, Number 12, issued 1989, N. Fotouhi et al., "Peptide Synthesis by Prior Thiol Capture. 6. Rates of the Disulfide Bond Forming of the Overall Strategy by Synthesis of the C-Terminal 29-Peptide Sequence of BPTI", pages 2803-2817.	1-27

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/07222

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	TETRAHEDRON LETT., Volume 32, Number 25, issued 1991, D.S. Kemp et al., "Resolution of the Histidine Problem for Thiol Capture-Synthesis of a 39-Peptide", pages 2845-2848.	1-27
A	J. ORG. CHEM., Volume 54, issued 1989, D.S. Kemp et al., "Rational Design of Templates for Intramolecular O,N-Acyl Transfer Via Medium-Sized Cyclic Intermediates Derived from L-Cysteine. Definition of an Experimental Maximum in Effective Molarity through the Study of 'Tunable' Templates", pages 1589-1603.	1-27
A	J. ORG. CHEM., Volume 51, issued 1986, D.S. Kemp et al., "Peptide Synthesis by Prior Thiol Capture. 4. Amide Bond Formation: The Effect of a Side-Chain Substituent on the Rates of Intramolecular O,N-Acyl Transfer", pages 3320-3324.	1-27, 43-48
A	INT. J. PEPTIDE PROTEIN RES., Volume 31, issued 1988, D. Yamashiro et al., "New Segment Synthesis of alpha-Inhibin-92 by the Acyl Disulfide Method", pages 322-334.	1-27

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/07222

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☒ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
1-27, 35-48

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/07222

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I, claims 1-16, drawn to a first method of ligating peptides, classified, for example, in Class 530, subclasses 333, 338, 339 and 402.

Group II, claims 17-27, drawn to second process, a method of making synthetic peptide dendrimers, classified, for example, in Class 436, subclass 543.

Group III, claims 28-34, drawn to first product, a peptide dendrimer prepared by the second process, classified, for example, in Class 530, subclasses 323, 403 and 807.

Group IV, claims 35-42, drawn to a third process, a method of modifying a protein with lipid, classified, for example, in Class 530, subclass 345.

Group V, claims 43-48, drawn to a fourth process, a method of site specific modification of a protein, classified, for example, in Class 530, subclass 402.

Group VI, claim 49, drawn to a second product, a lipid modified protein produced by the third process, classified, for example, in Class 530, subclass 359.

Group VII, claim 50, drawn to a third product, a specifically modified protein produced by the fourth process, classified, for example, in Class 530, subclass 402.

The inventions listed as Groups I-VII do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Groups I, II, IV and V are each drawn to distinctly different methods of using the basic conjugation chemistry of thiazolidine rearrangement. Group I is a method of combining two peptide groups, Group II is a method of attaching multiple copies of a desired peptide to several sites on the same polypeptide, Group IV is a method of attaching a aminoacyl-lipid to a protein, and Group V is a method introducing a thiazolidine group onto a protein. None of these methods requires all the particulars of the other groups and each of them results in a product with distinctly different structural properties. Each of these conjugations might be performed by other chemical reactions such as those cited in the accompanying references, e.g., enzyme catalyzed, anhydrides, or acyl imidazoles. Groups III, VI and VII are each distinct products with different structural properties and made by the distinctly different processes of Group II, IV and V respectively. Each of these products could be made by different chemistries as discussed above.
Because these inventions are distinct for the reasons given above and the Groups have acquired a separate status in the art as shown by their recognized divergent subject matter, lack of unity for examination purposes as indicated is proper.